

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2004, 07:44:20 ; Search time 18 Seconds

(without alignments)

108.988 Million cell updates/sec

Title: US-10-067-482-2_COPY_271_308

Perfect score: 200

Sequence: 1 GFLGLMGSGIVSNLLKMGHTVWNRVTAEKCDLFIQ 38

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Issued Patents AA:*

1: /cgn2_6/prodata/2/iaa/5A COMB.pep.*

2: /cgn2_6/prodata/2/iaa/5B COMB.pep.*

3: /cgn2_6/prodata/2/iaa/6A COMB.pep.*

4: /cgn2_6/prodata/2/iaa/6B COMB.pep.*

5: /cgn2_6/prodata/2/iaa/PCITUS COMB.pep.*

6: /cgn2_6/prodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	166	83.0	547	4	US-10-164-595-72
2	124	62.0	345	4	US-09-364-230-8
3	117	58.5	360	4	US-09-364-230-4
4	106	53.0	299	4	US-09-328-352-7407
5	102	51.0	326	4	US-09-328-352-8139
6	100	50.0	299	4	US-09-107-532A-5109
7	98	49.0	295	4	US-09-252-991A-24727
8	85	42.5	204	4	US-09-198-452A-376
9	84	42.0	304	4	US-09-489-039A-9424
10	81	40.5	247	4	US-09-364-230-2
11	80.5	40.2	307	4	US-09-489-039A-9294
12	77	38.5	330	4	US-09-252-991A-17923
13	75	37.5	221	4	US-09-252-991A-22003
14	75	37.5	474	4	US-09-134-001C-3241
15	74	37.0	490	4	US-09-489-039A-8350
16	73	36.5	474	3	US-09-058-692-2
17	73	36.5	474	3	US-09-584-628-2
18	72	36.0	486	4	US-09-543-681A-5380
19	69	34.5	324	4	US-09-252-991A-26001
20	68	34.0	74	4	US-09-621-976-6407
21	68	34.0	350	4	US-09-364-230-12
22	67	33.5	292	4	US-09-711-164-438
23	66	33.0	342	4	US-09-364-230-6
24	65	32.5	154	4	US-09-252-991A-29293
25	65	32.5	543	4	US-09-252-991A-27650
26	64	32.0	303	4	US-09-489-039A-12581
27	64	32.0	321	4	US-09-489-039A-10010

334	4	US-09-138-452A-923	Sequence 923, Ap
466	4	US-09-540-236-2874	Sequence 2874, Ap
352	4	US-09-134-001C-3137	Sequence 3137, Ap
361	30.5	US-09-107-532A-4004	Sequence 4004, Ap
32	59	US-09-634-955B-32	Sequence 32, Appl
33	59	US-09-489-039A-7231	Sequence 7231, Ap
34	59	US-09-328-352-6406	Sequence 6406, Ap
35	59	US-09-107-532A-7169	Sequence 7169, Ap
36	56	US-09-489-039A-7441	Sequence 7441, Ap
37	56	US-09-489-039A-13778	Sequence 13778, A
38	55	US-09-328-352-4879	Sequence 4879, Ap
39	55	US-08-745-934-4	Sequence 4, Appl
40	55	US-09-147-009-7	Sequence 7, Appl
54	54	US-09-489-039A-13223	Sequence 13223, A
42	54	US-09-134-000C-4723	Sequence 4723, Ap
43	53.5	US-09-252-991A-32168	Sequence 32168, A
44	53	US-08-220-602B-24	Sequence 24, Appl
45	53	US-09-107-532A-5053	Sequence 5053, Ap
46	53	US-09-252-991A-30567	Sequence 30567, A
47	53	US-08-311-731A-105	Sequence 105, Ap
48	52.5	US-09-415-277C-11	Sequence 11, Appl
49	52	US-09-540-236-3777	Sequence 3777, Ap
50	52	US-09-543-681A-4316	Sequence 4316, Ap
51	52	US-08-968-563-14	Sequence 14, Appl
52	52	US-08-969-683A-14	Sequence 14, Appl
53	52	US-09-297-928-10	Sequence 10, Appl
54	51.5	US-09-107-532A-7146	Sequence 7146, Ap
55	51	US-09-134-000C-4095	Sequence 4095, Ap
56	51	US-09-134-000C-5021	Sequence 5021, Ap
57	51	US-09-501-115-32	Sequence 32, Appl
58	51	US-09-489-039A-10351	Sequence 10351, A
59	51	US-09-489-039A-10528	Sequence 10528, A
60	51	US-09-252-991A-17150	Sequence 17150, A
61	51	US-08-946-329A-62	Sequence 62, Appl
62	50.5	US-09-328-352-6279	Sequence 6279, Ap
63	50.5	US-09-107-532A-7148	Sequence 7148, Ap
64	50.5	US-08-946-329A-17	Sequence 17, Appl
65	50.5	US-08-567-357A-17	Sequence 17, Appl
66	50.5	US-08-729-743A-17	Sequence 17, Appl
67	50.5	US-08-349-498-17	Sequence 17, Appl
68	50.5	US-09-325-256-25	Sequence 25, Appl
69	50.5	US-09-704-917-18	Sequence 18, Appl
70	50.5	US-09-151-999-18	Sequence 18, Appl
71	50.5	PCT-US95-15463-17	Sequence 17, Appl
72	50.5	PCT-US95-15923-17	Sequence 17, Appl
73	50.5	US-09-134-001C-3055	Sequence 3055, Ap
74	50.5	US-09-149-476-440	Sequence 440, Ap
75	50	US-09-543-681A-7202	Sequence 7202, Ap
76	50	US-09-231-227-4	Sequence 4, Appl
77	50	US-09-768-585-4	Sequence 4, Appl
78	50	US-09-489-039A-13474	Sequence 13474, A
79	50	US-09-134-001C-5237	Sequence 5237, Ap
80	50	US-09-252-991A-33108	Sequence 33108, A
81	50	US-09-134-001C-5065	Sequence 11014, A
82	50	US-09-489-039A-11014	Sequence 6726, Ap
83	50	US-09-328-352-6726	Sequence 1, Appl
84	50	US-08-745-934-1	Sequence 29547, A
85	50	US-09-252-991A-29547	Sequence 10777, A
86	49.5	US-09-058-333A-2	Sequence 2, Appl
87	49.5	US-09-489-039A-10777	Sequence 2, Appl
88	49.5	US-08-199-737-2	Sequence 2, Appl
89	49.5	US-08-900-230-2	Sequence 2, Appl
90	49.5	US-09-058-333A-2	Sequence 2, Appl
91	49.5	US-09-711-164-335	Sequence 335, Ap
92	49	US-09-540-236-3409	Sequence 3409, Ap
93	49	US-08-931-668-9	Sequence 9, Appl
94	49	US-09-468-175-9	Sequence 9, Appl
95	49	US-09-154-874-8	Sequence 8, Appl
96	49	US-08-931-668-8	Sequence 8, Appl
97	49	US-09-468-175-8	Sequence 8, Appl
98	49	US-09-543-681A-7503	Sequence 7503, Ap
99	49	US-09-489-039A-13595	Sequence 13595, A
100	49		

APPLICANT: Kinney, Anthony J.
APPLICANT: Rafalski, J. Antoni

Db 49 FLGIGLMSRMSRLIQAGFQVAVMNRITSACEELI 84

RESULT 6

US-09-107-532A-5109

; Sequence 5109, Application US/09107532A

; Patent No. 6583275

; GENERAL INFORMATION:

; APPLICANT: Lynn A Doucette-Stamm and David Bush

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

; NUMBER OF SEQUENCES: 7310

; CORRESPONDENCE ADDRESSES:

; ADDRESSEE: GENOME THERAPEUTICS CORPORATION

; STREET: 100 Beaver Street

; CITY: Waltham

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02354

; COMPUTER READABLE FORM:

; MEDIUM TYPE: CD/ROM ISO9660

; COMPUTER: PC

; OPERATING SYSTEM: <Unknown>

; SOFTWARE: ASCII

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/107,532A

; FILING DATE: 30-Jun-1998

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/085,598

; FILING DATE: 14 May 1998

; APPLICATION NUMBER: 60/051571

; FILING DATE: July 2, 1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Ariniello, Pamela Deneke

; REGISTRATION NUMBER: 40,489

; REFERENCE/DOCKET NUMBER: GTC-012

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (781)893-5007

; TELEFAX: (781)893-8277

; INFORMATION FOR SEQ ID NO: 5109:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 299 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHETICAL: YES

; ORIGINAL SOURCE:

; ORGANISM: Enterococcus faecium

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (B) LOCATION 1...299

; SEQUENCE DESCRIPTION: SEQ ID NO: 5109:

US-09-107-532A-5109

Query Match 50.0%; Score 100; DB 4; Length 299;

Best Local Similarity 52.9%; Pred. No. 2.6e-06;

Matches 18; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 1 GFGLGLMGSGIVSNLLKMGHTVTVMNRITAEKCD 34

Db 8 GFIGTGWGSARHLEAGHEGAVVNRITKAD 41

RESULT 7

US-09-252-991A-24727

; Sequence 24727, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

RESULT 12
US-09-252-991A-17923
; Sequence 17923, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17923
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17923
Query Match 38.5%; Score 77; DB 4; Length 330;
Best Local Similarity 47.1%; Pred. No. 0.0054;
Matches 16; Conservative 5; Mismatches 13; Indels 0; Gaps 0;
Qy 2 FLGLMGSGIVSNLLKMGHTVTVMNRTAEKCDL 35
Db 44 FAGLGMGVPICRELLAAGYFLAVNRSPOKREL 77
RESULT 13
US-09-252-991A-22003
; Sequence 22003, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22003
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22003
Query Match 37.5%; Score 75; DB 4; Length 221;
Best Local Similarity 57.7%; Pred. No. 0.0065;
Matches 15; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
Qy 2 FLGLMGSGIVSNLLKMGHTVTVMN 27
Db 11 FLGLNMGPMANLLKAGHRVNVFD 36
RESULT 14
US-09-134-001C-3241
; Sequence 3241, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13

Qy 3 LGLGLMGSGIVSNLLKMGHTVTVMNRT 29
Db 23 LGLGAMGHAFASNLLKMGFTVAGWNR 49
RESULT 10
US-09-364-230-2
; Sequence 2, Application US/09364230
; Patent No. 6348339
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Hitz, William D.
; APPLICANT: Kinney, Anthony J.
; APPLICANT: Rafalski, J. Antoni
; TITLE OF INVENTION: Enzymes Involved in Degradation of Branched-Chain Amino Acids
; FILE REFERENCE: BB-1178
; CURRENT APPLICATION NUMBER: US/09/364,230
; CURRENT FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,990
; EARLIER FILING DATE: July 31, 1998
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (244)
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (247)
US-09-364-230-2
Query Match 40.5%; Score 81; DB 4; Length 247;
Best Local Similarity 71.4%; Pred. No. 0.001;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
Qy 14 SNLLKMGHTVTVMNRTAEKCD 34
Db 2 SNLLKAGCDVTVMNRTSKCD 22
RESULT 11
US-09-489-039A-9294
; Sequence 9294, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9294
; LENGTH: 307
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9294
Query Match 40.2%; Score 80.5; DB 4; Length 307;
Best Local Similarity 31.0%; Pred. No. 0.0016;
Matches 18; Conservative 10; Mismatches 9; Indels 21; Gaps 1;
Qy 1 GFLGLMGSGIVSNLLKMGHTVTVMNRT-----AEKCDLFI 37
Db 17 GFGLGIMGKPMKSNLLKAGYSLVSDRNPEAIDVIAAGATATTPKAIABQCEVII 74

```

; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Falk, Stephen T
; REGISTRATION NUMBER: 36,795
; REFERENCE/DOCKET NUMBER: GM10140
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2488
; TELEFAX: 215-994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 474 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-058-692-2

Query Match 36.5%; Score 73; DB 3; Length 474;
Best Local Similarity 35.3%; Pred.No. 0.031;
Matches 12; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

QY      1  GFGLGIMGSGIYSNLLKMGHTVTVMRTAEKCD 34
        |+:|::||: |:|::||:|::||:|:
Db       7  GVGMVAMGRNLALNIESRGYVAIYNRSKEKT 40

RESULT 17
US-09-584-628-2
; Sequence 2, Application US/09584628
; Patent No. 6309866
; GENERAL INFORMATION:
; APPLICANT: Warren, Richard L.
; TITLE OF INVENTION: 6-phosphogluconate dehydr
; TITLE OF INVENTION: ogenase
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert, Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2793
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/584,628
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/058,692
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Falk, Stephen T
; REGISTRATION NUMBER: 36,795
; REFERENCE/DOCKET NUMBER: GM10140
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2488
; TELEFAX: 215-994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 474 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-584-628-2

```

```
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 6407
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-6407

Query Match          34.0%; Score 68; DB 4; Length 74;
Best Local Similarity 37.1%; Pred.No. 0.017;
Matches 13; Conservative 7; Mismatches 15; Indels 0; Gaps 0;

QY      3 IGLGLMGSGIVSNLLKMGHTVTVWNRATAEKCDLFI 37
       :|||:::||: |: |:: ||| |:: ||| :
Ddb     9 IGLAVMGNQLILNMNDHGFEVCAFNRTVSKVDDEL 43

RESULT 21
US-09-364-230-12
; Sequence 12, Application US/09364230
; Patent No. 6348339
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Hitz, William D.
; APPLICANT: Kinney, Anthony J.
; APPLICANT: Rafalski, J. Antoni
; TITLE OF INVENTION: Enzymes Involved in Degradation of Branched-Chain Amino Acids
; FILE REFERENCE: BB-1178
; CURRENT APPLICATION NUMBER: US/09/364,230
; CURRENT FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,990
; EARLIER FILING DATE: July 31, 1998
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (154)
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (176)
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (184)..(201)
US-09-364-230-12

Query Match          34.0%; Score 68; DB 4; Length 350;
Best Local Similarity 52.0%; Pred.No. 0.11;
Matches 13; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY      1 GFGLGLMGSGIVSNLLKMGTITV 25
       |||::|||:: |||:: |||:: ||| :
Ddb     44 GFIGLGNGSMHARNLVAGRYRVSF 68

RESULT 22
US-09-711-164-438
; Sequence 438, Application US/09711164
; Patent No. 6589738
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Orlsen, Kari
```

```
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29293

Query Match      32.5%; Score 65; DB 4; Length 154;
Best Local Similarity 46.2%; Pred. No. 0.11;
Matches 12; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

QY       2 FLGLMGSGIVSNLLKMGHTVTVWN 27
         |::|::|::|::|::|::|::|::|:|::|
Db        43 FIGLHGAGPATNLLKAGYLNVFD 68

RESULT 25
US-09-252-991A-27650
; Sequence 27650, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27650
; LENGTH: 543
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27650

Query Match      32.5%; Score 65; DB 4; Length 543;
Best Local Similarity 47.8%; Pred. No. 0.49;
Matches 11; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY       1 GFGLGMLGMGSIGVSNLLKMGTW 23
         ||::||::|::|::|::|::|::|::|
Db        252 GRFGTGIMCKPVAQNLOKAGHSL 274

RESULT 26
US-09-489-039A-12581
; Sequence 12581, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONTAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12581
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12581

Query Match      32.0%; Score 64; DB 4; Length 303;
Best Local Similarity 48.0%; Pred. No. 0.34;
Matches 12; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY       4 GLGLMGSGIVSNLLKMGHTVTVNR 28
         ||::|::|::|::|::|::|::|::|
Db        14 GTGLGRHLIPFLLELGHSVTVSTR 38
```

```

; ORGANISM: M.catarrhalis
US-09-540-236-2874

Query Match      31.0%; Score 62; DB 4; Length 466;
Best Local Similarity 42.3%; Pred. No. 1.1;
Matches 11; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-03-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10010
; LENGTH: 321
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10010

Query Match      32.0%; Score 64; DB 4; Length 321;
Best Local Similarity 34.4%; Pred. No. 0.37;
Matches 11; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3137
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3137

Query Match      30.5%; Score 61; DB 4; Length 352;
Best Local Similarity 32.3%; Pred. No. 1.1;
Matches 10; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2874
; LENGTH: 466
; TYPE: PRT

```

```

; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-03-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10010
; LENGTH: 321
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10010

Query Match      32.0%; Score 64; DB 4; Length 321;
Best Local Similarity 34.4%; Pred. No. 0.37;
Matches 11; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2874
; LENGTH: 466
; TYPE: PRT

```

```

RESULT 27
US-09-489-039A-10010
; Sequence 10010, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-03-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10010
; LENGTH: 321
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10010

Query Match      32.0%; Score 64; DB 4; Length 321;
Best Local Similarity 34.4%; Pred. No. 0.37;
Matches 11; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2874
; LENGTH: 466
; TYPE: PRT

```

```

RESULT 28
US-09-198-452A-923
; Sequence 923, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffiths, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 923
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-923

Query Match      31.5%; Score 63; DB 4; Length 334;
Best Local Similarity 35.5%; Pred. No. 0.53;
Matches 11; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2874
; LENGTH: 466
; TYPE: PRT

```

```

RESULT 29
US-09-540-236-2874
; Sequence 2874, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2874
; LENGTH: 466
; TYPE: PRT

```

```

; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2874
; LENGTH: 466
; TYPE: PRT

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2004, 07:45:18 ; Search time 48 Seconds
(without alignments)
254.231 Million cell updates/sec

Title: US-10-067-482-2_COPY_271_308

Perfect score: 200

Sequence: 1 GFGLGMSGIVSNLLKMGHTVTVMNRTAEKCDLFIQ 38

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1342398 seqs, 321133274 residues

Total number of hits satisfying chosen parameters: 1342398

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	200	100.0	553	14	US-10-067-482-2
2	166	83.0	547	14	US-10-067-482-3
3	166	83.0	550	14	US-10-103-313-417
4	164	82.0	276	14	US-10-067-482-4
5	125	62.5	289	14	US-10-167-547C-16
6	124	62.0	333	12	US-10-424-599-144794
7	121	60.5	176	16	US-10-767-701-60840
8	121	60.5	343	16	US-10-437-963-116017
9	118	59.0	290	14	US-10-167-547C-14
10	118	59.0	293	16	US-10-437-963-174476
11	117	58.5	81	16	US-10-767-701-52389
12	117	58.5	364	12	US-10-425-114-50561
13	109	54.5	292	12	US-10-282-122A-52083
14	106	53.0	288	12	US-10-282-122A-45255
15	105	52.5	290	12	US-10-282-122A-44990
16	105	52.5	290	12	US-10-282-122A-44990
17	105	52.5	290	12	US-10-282-122A-44990
18	105	52.5	290	12	US-10-282-122A-44990
19	105	52.5	290	12	US-10-282-122A-44990
20	105	52.5	290	12	US-10-282-122A-44990
21	105	52.5	290	12	US-10-282-122A-44990
22	105	52.5	290	12	US-10-282-122A-44990
23	105	52.5	290	12	US-10-282-122A-44990
24	105	52.5	290	12	US-10-282-122A-44990
25	105	52.5	290	12	US-10-282-122A-44990
26	105	52.5	290	12	US-10-282-122A-44990
27	105	52.5	290	12	US-10-282-122A-44990
28	105	52.5	290	12	US-10-282-122A-44990
29	105	52.5	290	12	US-10-282-122A-44990
30	105	52.5	290	12	US-10-282-122A-44990
31	105	52.5	290	12	US-10-282-122A-44990
32	105	52.5	290	12	US-10-282-122A-44990
33	105	52.5	290	12	US-10-282-122A-44990
34	105	52.5	290	12	US-10-282-122A-44990
35	105	52.5	290	12	US-10-282-122A-44990
36	105	52.5	290	12	US-10-282-122A-44990
37	105	52.5	290	12	US-10-282-122A-44990
38	105	52.5	290	12	US-10-282-122A-44990
39	105	52.5	290	12	US-10-282-122A-44990
40	105	52.5	290	12	US-10-282-122A-44990
41	105	52.5	290	12	US-10-282-122A-44990
42	105	52.5	290	12	US-10-282-122A-44990
43	105	52.5	290	12	US-10-282-122A-44990
44	105	52.5	290	12	US-10-282-122A-44990
45	105	52.5	290	12	US-10-282-122A-44990
46	105	52.5	290	12	US-10-282-122A-44990
47	105	52.5	290	12	US-10-282-122A-44990
48	105	52.5	290	12	US-10-282-122A-44990
49	105	52.5	290	12	US-10-282-122A-44990
50	105	52.5	290	12	US-10-282-122A-44990
51	105	52.5	290	12	US-10-282-122A-44990
52	105	52.5	290	12	US-10-282-122A-44990
53	105	52.5	290	12	US-10-282-122A-44990
54	105	52.5	290	12	US-10-282-122A-44990
55	105	52.5	290	12	US-10-282-122A-44990
56	105	52.5	290	12	US-10-282-122A-44990
57	105	52.5	290	12	US-10-282-122A-44990
58	105	52.5	290	12	US-10-282-122A-44990
59	105	52.5	290	12	US-10-282-122A-44990
60	105	52.5	290	12	US-10-282-122A-44990
61	105	52.5	290	12	US-10-282-122A-44990
62	105	52.5	290	12	US-10-282-122A-44990
63	105	52.5	290	12	US-10-282-122A-44990
64	105	52.5	290	12	US-10-282-122A-44990
65	105	52.5	290	12	US-10-282-122A-44990
66	105	52.5	290	12	US-10-282-122A-44990
67	105	52.5	290	12	US-10-282-122A-44990
68	105	52.5	290	12	US-10-282-122A-44990
69	105	52.5	290	12	US-10-282-122A-44990
70	105	52.5	290	12	US-10-282-122A-44990
71	105	52.5	290	12	US-10-282-122A-44990
72	105	52.5	290	12	US-10-282-122A-44990
73	105	52.5	290	12	US-10-282-122A-44990
74	105	52.5	290	12	US-10-282-122A-44990
75	105	52.5	290	12	US-10-282-122A-44990
76	105	52.5	290	12	US-10-282-122A-44990
77	105	52.5	290	12	US-10-282-122A-44990
78	105	52.5	290	12	US-10-282-122A-44990
79	105	52.5	290	12	US-10-282-122A-44990
80	105	52.5	290	12	US-10-282-122A-44990
81	105	52.5	290	12	US-10-282-122A-44990
82	105	52.5	290	12	US-10-282-122A-44990
83	105	52.5	290	12	US-10-282-122A-44990
84	105	52.5	290	12	US-10-282-122A-44990
85	105	52.5	290	12	US-10-282-122A-44990
86	105	52.5	290	12	US-10-282-122A-44990
87	105	52.5	290	12	US-10-282-122A-44990
88	105	52.5	290	12	US-10-282-122A-44990

Sequence 57891, A
Sequence 11979, A
Sequence 9081, App
Sequence 56894, A
Sequence 47190, A
Sequence 5781, App
Sequence 12590, A
Sequence 2943, App
Sequence 258895, A
Sequence 43907, A
Sequence 3, Appli
Sequence 3, Appli
Sequence 60359, A
Sequence 5105, App
Sequence 61555, A
Sequence 376, App
Sequence 146727, A
Sequence 54804, A
Sequence 55145, A
Sequence 53968, A
Sequence 256, App
Sequence 76475, A
Sequence 150481, A
Sequence 178316, A
Sequence 58955, A
Sequence 13226, A
Sequence 23377, A
Sequence 8219, App
Sequence 255872, A
Sequence 255869, A
Sequence 4866, App
Sequence 7628, App
Sequence 49776, A
Sequence 163370, A
Sequence 60763, A
Sequence 11686, A
Sequence 15834, A
Sequence 62538, A
Sequence 19951, A
Sequence 19011, A
Sequence 11054, A
Sequence 58219, A
Sequence 208, App
Sequence 64615, A
Sequence 22002, A
Sequence 301, App
Sequence 20250, A
Sequence 71412, A
Sequence 67336, A
Sequence 22062, A
Sequence 209573, A
Sequence 208870, A
Sequence 42795, A
Sequence 19494, A
Sequence 219746, A
Sequence 44516, A
Sequence 203136, A
Sequence 202762, A
Sequence 47392, A
Sequence 50082, A
Sequence 201442, A
Sequence 53209, A
Sequence 59273, A
Sequence 46495, A
Sequence 46975, A
Sequence 47502, A
Sequence 47600, A
Sequence 53351, A
Sequence 59175, A
Sequence 59189, A
Sequence 59993, A
Sequence 62688, A
Sequence 66115, A

```

89      77      38.5      485      12      US-10-425-114-72073      Sequence 72073, A
90      77      38.5      485      12      US-10-425-114-72092      Sequence 72092, A
91      77      38.5      485      12      US-10-425-114-72099      Sequence 72099, A
92      76      38.0      298      15      US-10-369-493-8964      Sequence 8964, Ap
93      76      38.0      468      9       US-09-875-573-43       Sequence 43, Appl
94      76      38.0      471      15      US-10-369-493-20932      Sequence 20932, A
95      76      38.0      483      12      US-10-282-122A-77672      Sequence 77672, A
96      76      38.0      483      12      US-10-282-122A-63994      Sequence 63994, A
97      76      38.0      489      12      US-10-282-122A-61885      Sequence 61885, A
98      75      37.5      274      16      US-10-767-701-41225      Sequence 41225, A
99      75      37.5      468      12      US-10-282-122A-70621      Sequence 70621, A
100     75      37.5      495      12      US-10-425-114-65181      Sequence 65181, A

ALIGNMENTS

RESULT 1
US-10-067-482-2
; Sequence 2, Application US/10067482
; Publication No. US20030148407A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc.
; TITLE OF INVENTION: Angiogenesis Dehydrogenase Gene
; FILE REFERENCE: IU 102 RI
; CURRENT APPLICATION NUMBER: US/10/067,482
; CURRENT FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 553
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-067-482-2
Query Match      100.0%; Score 200; DB 14; Length 553;
Best Local Similarity 100.0%; Pred. No. 2.7e-19;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1      GFLGLGMSGIVSNLLKMGHTVTVWNRRTAEK 38
Db      271      GFLGLGMSGIVSNLLKMGHTVTVWNRRTAEK 308

RESULT 2
US-10-067-482-3
; Sequence 3, Application US/10067482
; Publication No. US20030148407A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc.
; TITLE OF INVENTION: Angiogenesis Dehydrogenase Gene
; FILE REFERENCE: IU 102 RI
; CURRENT APPLICATION NUMBER: US/10/067,482
; CURRENT FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 547
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-067-482-3
Query Match      83.0%; Score 166; DB 14; Length 547;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1      GFLGLGMSGIVSNLLKMGHTVTVWNRRTAEK 32
Db      271      GFLGLGMSGIVSNLLKMGHTVTVWNRRTAEK 302

RESULT 3
US-10-103-313-417
; Sequence 417, Application US/10103313
; Publication No. US20030082758A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ207C1
; CURRENT APPLICATION NUMBER: US/10/103,313
; CURRENT FILING DATE: 2002-03-12
; NUMBER OF SEQ ID NOS: 653
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 417
; LENGTH: 550
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-103-313-417
Query Match      83.0%; Score 166; DB 14; Length 550;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1      GFLGLGMSGIVSNLLKMGHTVTVWNRRTAEK 32
Db      274      GFLGLGMSGIVSNLLKMGHTVTVWNRRTAEK 305

RESULT 4
US-10-067-482-4
; Sequence 4, Application US/10067482
; Publication No. US20030148407A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc.
; TITLE OF INVENTION: Angiogenesis Dehydrogenase Gene
; FILE REFERENCE: IU 102 RI
; CURRENT APPLICATION NUMBER: US/10/067,482
; CURRENT FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 276
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-067-482-4
Query Match      82.0%; Score 164; DB 14; Length 276;
Best Local Similarity 100.0%; Pred. No. 1.3e-14;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      8      MSGIVSNLLKMGHTVTVWNRRTAEKCDLFIQ 38
Db      1      MSGIVSNLLKMGHTVTVWNRRTAEKCDLFIQ 31

RESULT 5
US-10-167-547C-16
; Sequence 16, Application US/10167547C
; Publication No. US20030170653A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont de Nemours and Company
; APPLICANT: Damude, Howard G.
; TITLE OF INVENTION: A Biological Method for the Production of Alpha-Methylene-Gamma
; FILE REFERENCE: CL1804 US NA
; CURRENT APPLICATION NUMBER: US/10/167,547C
; CURRENT FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: 60/297198
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Microsoft Office 07
; SEQ ID NO 16
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Arabidopsis
```



```

Query Match          60.5%;   Score 121;   DB 16;   Length 176;
Best Local Similarity 64.7%;   Pred. No. 8.5e-09;
Matches 22;   Conservative 4;   Mismatch 8;   Indels 0;   Gaps 0;

1  GFLGGLMGSGIVSNLLKNGHTVTVMNRTEKCD 34
67  GFLGGLMGAPASNLISAGCDVTVMNRTEKCD 100

```

RESULT 10
US-10-437-963-174476
; Sequence 174476, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:

```

; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 174476
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_72412C.1.pep
; US-10-437-963-174476

Query Match      59.8%; Score 118; DB 16; Length 293;
Best Local Similarity 63.6%; Pred. No. 4e-08;
Matches 21; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY      1  GFLGLGMLGSGIVSNLLKMGHTVTVVNRRTAEKC 33
        |||||:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      4  GFLGLGIMGKAMAANLLRHGFRVTVVNRRTLAKC 36

RESULT 11
US-10-767-701-52389
; Sequence 52389, Application US/10767701
; Publication No. US20040175684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 52389
; LENGTH: 81
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 11922868.pep
; US-10-767-701-52389

Query Match      58.5%; Score 117; DB 16; Length 81;
Best Local Similarity 63.6%; Pred. No. 1.3e-08;
Matches 21; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY      1  GFLGLGMLGSGIVSNLLKMGHTVTVVNRRTAEKC 33
        |||||:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      4  GFLGLGIMGKAMAANLLRHGFRVTVVNRRTLAKC 36

RESULT 12
US-10-425-114-50561
; Sequence 50561, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 50561
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3136-019-H11_FLI.pep
; US-10-425-114-50561

Query Match      58.5%; Score 117; DB 12; Length 364;
Best Local Similarity 63.6%; Pred. No. 7e-08;
Matches 21; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY      1  GFLGLGMLGSGIVSNLLKMGHTVTVVNRRTAEKC 33
        |||||:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      73 GFLGLGIMGKAMATNLLRHGFRVTVVNRRTLAKC 105

RESULT 13
US-10-282-122A-52083
; Sequence 52083, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 52083
; LENGTH: 292
; TYPE: PRT
; ORGANISM: Clostridium acetobutylicum
; US-10-282-122A-52083
```

```

Query Match          54.5%;   Score 109;   DB 12;   Length 292;
Best Local Similarity 62.5%;   Pred. No. 7.3e-07;
Matches 20; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 1 FLGLGLMGSGIVSNLLKMGHTVTVWNRATK 32
DB 9 FLGLGLMGSGIVSNLLKMGHTVTVWNRATK 40

RESULT 14
US-10-282-122A-45255
; Sequence 45255, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 45255
; LENGTH: 288
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-10-282-122A-45255

Query Match          53.0%;   Score 106;   DB 12;   Length 288;
Best Local Similarity 55.6%;   Pred. No. 1.9e-06;
Matches 20; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 2 FLGLGLMGSGIVSNLLKMGHTVTVWNRATKCDLFI 37
DB 11 FLGLGLMGSRMASRLIQAGFQVAVWNRATASACEELI 46

RESULT 15
US-10-282-122A-44990
; Sequence 44990, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:

```

```

RESULT 18
US-10-156-761-9081
; Sequence 9081, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIXAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9081
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9081

Query Match          46.5%;   Score 93;   DB 14;   Length 296;
Best Local Similarity 50.0%;   Pred. No. 0.00013;
Matches 16; Conservative 5; Mismatches 11; Indels 0; Gaps 0

Qy      3 LGLGIMGSGIVSNLLKMGHTVTVMNRTAEKCD 34
        ||||| : || : |||| |||| | :
Db      10 LGTGIMGAAMARNLARAGHTVRAWNRTRAKAE 41

RESULT 19
US-10-282-122A-56894
; Sequence 56894, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22

```

```

Best Local Similarity 42.1%; Pred. No. 0.0011;
Matches 16; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

QY 1 GFLGLGLMGSGIVSNLLKMGHTVTVMNRTAEKCDLFIQ 38
DB 4 GYGLGVWGSNLIANIADNGFVSVFNRDNEKTEIFVK 41

RESULT 21
US-09-815-242-5781
; Sequence 5781, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5781
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5781

Query Match 44.0%; Score 88; DB 9; Length 465;
Best Local Similarity 39.5%; Pred. No. 0.0011;
Matches 15; Conservative 12; Mismatches 11; Indels 0; Gaps 0;

QY 1 GFLGLGLMGSGIVSNLLKMGHTVTVMNRTAEKCDLFIQ 38
DB 4 GVIGLAVMGKNLAWNIEGRGYSVSVFNRSEKTDLMVE 41

RESULT 22
US-09-815-242-12590
; Sequence 12590, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A

```

```

; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12590
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12590

Query Match          44.0%; Score 88; DB 9; Length 468;
Best Local Similarity 39.5%; Pred. No. 0.0011;
Matches 15; Conservative 12; Mismatches 11; Indels 0; Gaps 0;

QY 1 GFLGLGMSGIVSNLLKMGHTVTWNRATAKCDLFIQ 38
Db 6 GVIGLAVMGKLNIAWNIERSGYSVFNRSSEKTLME 43

RESULT 23
US-10-369-493-2943
; Sequence 2943, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; CURRENT APPLICATION NUMBER: US/10/369,493
; FILE REFERENCE: 38-10(52052)B
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 2943
; LENGTH: 469
; TYPE: PRT
; ORGANISM: Thermotoga maritima
US-10-369-493-2943

Query Match          44.0%; Score 88; DB 15; Length 469;
Best Local Similarity 36.8%; Pred. No. 0.0011;
Matches 14; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

QY 1 GFLGLGMSGIVSNLLKMGHTVTWNRATAKCDLFIQ 38
Db 6 GLIGLAVMGONLAINIARKYKYSVYNRTAQTDFEVK 43

RESULT 24
US-10-424-599-258895
; Sequence 258895, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 258895
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_75806C.1.pep
US-10-424-599-258895

Query Match          43.5%; Score 87; DB 12; Length 122;
Best Local Similarity 40.5%; Pred. No. 0.00033;
Matches 15; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

QY 1 GFLGLGMSGIVSNLLKMGHTVTWNRATAKCDLFI 37
Db 22 GLIGLGVMSNLAINIAEKGNRIAVFNRTPARTDFEL 58

RESULT 25
US-10-282-122A-43907
; Sequence 43907, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43907
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
```


Thu Sep 16 09:36:50 2004

```

; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 5105
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5105

Query Match          43.0%; Score 86; DB 9; Length 492;
Best Local Similarity 43.2%; Pred. No. 0.0022;
Matches 16; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

QY 1 GFLGLGMSGIVSNLLKMGHTVTVMNRTAEKCDLFI 37
DB 19 GVGGLAVGWSNLARNFARNGTAVVNRSTDKTKLI 55

RESULT 30
US-10-767-701-61555
; Sequence 61555, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 61555
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 9847593.pep
US-10-767-701-61555

Query Match          42.5%; Score 85; DB 16; Length 154;
Best Local Similarity 40.5%; Pred. No. 0.00081;
Matches 15; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

QY 1 GFLGLGMSGIVSNLLKMGHTVTVMNRTAEKCDLFI 37
DB 26 GWIGIGVGGGAKRLLAAGYTVTAYARTPSKAGFV 62

Search completed: September 16, 2004, 07:47:49
Job time : 51 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2004, 07:32:13 ; Search time 15 Seconds

(without alignments)

243.685 Million cell updates/sec

Title: US-10-067-482-2_COPY_271_308

Perfect score: 200

Sequence: 1 GFLGLMGSGIVSNLLKMGHTVWNRPAEKDGLFIQ 38

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	109	54.5	292	2	G97310 dehydrogenase rela
2	108	54.0	289	2	G90314 oxidoreductase [im
3	99	49.5	288	2	B69870 3-hydroxyisobutyra
4	98	49.0	291	2	D83371 probable dehydroge
5	92	46.0	334	2	T08967 hypothetical prote
6	90	45.0	293	2	G95277 probable D-threoni
7	90	45.0	470	1	S14628 phosphogluconate d
8	88	44.0	464	2	H70169 phosphogluconate d
9	88	44.0	468	2	G89930 6-phosphogluconate
10	88	44.0	469	2	A72377 6-phosphogluconate d
11	86	43.0	286	2	ACL558 3-hydroxyisobutyra
12	86	43.0	472	2	AD1609 6-phosphogluconate
13	86	43.0	472	2	AH1246 6-phosphogluconate
14	85	42.5	479	2	F86535 6-phosphogluconate
15	85	42.5	479	2	C72088 6-phosphogluconate
16	85	42.5	479	2	A81714 6-phosphogluconate
17	84.5	42.2	299	1	U00613 3-hydroxyisobutyra
18	84.5	42.2	299	2	C91129 probable dehydroge
19	84.5	42.2	299	2	C85974 probable dehydroge
20	84	42.0	288	2	C70303 3-hydroxyisobutyra
21	84	42.0	295	2	A96019 probable dehydroge
22	84	42.0	299	2	D96736 probable dehydroge
23	84	42.0	488	2	A71337 probable phosphogl
24	83	41.5	468	1	D26190 phosphogluconate d
25	82.5	41.2	291	2	AB3380 3-hydroxyisobutyra
26	82.5	41.2	294	2	AE0897 2-hydroxy-3-oxopro
27	82	41.0	293	2	G86901 hypothetical prote
28	82	41.0	297	2	E96736 probable dehydroge
29	82	41.0	300	2	AH2912 oxidoreductase [imp

30	82	41.0	300	2	C97687 D-threonine dehydr
31	81	40.5	286	2	AE1200 3-hydroxyisobutyra
32	81	40.5	467	2	JC2306 phosphogluconate d
33	81	40.5	476	2	AE2764 phosphogluconate d
34	81	40.5	476	2	D97545 6-phosphogluconate
35	80	40.0	289	2	H90248 3-hydroxyisobutyra
36	80	40.0	298	2	E91229 probable dehydroge
37	80	40.0	288	2	D86076 probable dehydroge
38	80	40.0	298	2	S40826 hypothetical 31.2K
39	80	40.0	298	2	AI0947 probable oxidoredu
40	80	40.0	476	2	AC2465 6-phosphogluconate
41	80	40.0	479	2	A48565 phosphogluconate d
42	80	40.0	484	2	C64077 phosphogluconate d
43	80	40.0	485	2	D70864 phosphogluconate d
44	80	40.0	492	2	S64588 probable gnd prote
45	79	39.5	469	2	AC3650 phosphogluconate d
46	79	39.5	480	2	A71561 phosphogluconate d
47	79	39.5	483	2	JE0234 probable 6-phospho
48	79	39.5	489	1	S46671 phosphogluconate d
49	77	38.5	296	2	C83232 phosphogluconate d
50	77	38.5	315	2	H82512 probable 3-hydroxy
51	77	38.5	482	2	T01659 phosphogluconate d
52	77	38.5	484	2	JC5282 phosphogluconate d
53	77	38.5	484	2	T01658 phosphogluconate d
54	76	38.0	226	2	S76263 phosphogluconate d
55	76	38.0	286	2	E87444 hypothetical prote
56	76	38.0	482	2	C82404 conserved hypothet
57	76	38.0	483	2	D87167 6-phosphogluconate
58	76	38.0	486	2	T44750 probable phosphogl
59	75	37.5	298	1	C42902 3-hydroxyisobutyra
60	75	37.5	371	2	T10635 3-hydroxyisobutyra
61	75	37.5	482	2	E81248 6-phosphogluconate
62	74	37.0	125	2	H47677 hypothetical prote
63	74	37.0	296	2	T34859 probable dehydroge
64	74	37.0	315	2	A98338 D-threonin dehydro
65	74	37.0	315	2	AI2944 3-hydroxyisobutyra
66	74	37.0	468	1	S04397 phosphogluconate d
67	74	37.0	468	2	I62463 phosphogluconate d
68	74	37.0	468	2	D56146 phosphogluconate d
69	74	37.0	468	2	I84555 phosphogluconate d
70	74	37.0	468	2	I41250 phosphogluconate d
71	74	37.0	468	2	AE0765 6-phosphogluconate
72	74	37.0	492	2	T40628 6-phosphogluconate
73	74	37.0	508	2	T42523 probable phosphogl
74	74	37.0	511	2	T05363 phosphogluconate d
75	73	36.5	290	2	E72666 probable 3-hydroxy
76	73	36.5	287	2	S76381 probable 3-hydroxy
77	73	36.5	443	2	I40708 phosphogluconate d
78	73	36.5	468	1	DEECGC phosphogluconate d
79	73	36.5	468	2	I62465 phosphogluconate d
80	73	36.5	468	2	I41249 phosphogluconate d
81	73	36.5	468	2	F90982 gluconate-6-phosph
82	73	36.5	468	2	D85828 gluconate-6-phosph
83	73	36.5	472	2	F86702 gluconate-6-phosph
84	73	36.5	474	2	E95043 hypothetical prote
85	73	36.5	474	2	G97913 hypothetical prote
86	73	36.5	482	2	B82021 phosphogluconate d
87	72	36.0	211	2	AC2560 phosphogluconate d
88	72	36.0	469	2	AI0187 hypothetical prote
89	72	36.0	486	2	S57786 phosphogluconate d
90	71	35.5	356	2	E71252 phosphogluconate d
91	70	35.0	292	2	AB0567 probable glycerol -
92	70	35.0	298	2	C96022 2-hydroxy-3-oxopro
93	70	35.0	310	2	AB3338 probable 3-hydroxy
94	69	34.5	285	2	G71912 3-hydroxyisobutyra
95	69	34.5	288	2	F83447 probable 3-hydroxy
96	69	34.5	297	2	D91027 probable 3-hydroxy
97	69	34.5	297	2	E85871 probable sugar nuc
98	69	34.5	346	2	A32867 3-hydroxyisobutyra
99	68	34.0	289	2	B81065 3-hydroxyacid dehy
100	68	34.0	289	2	B81802 hypothetical prote

RESULT 5
T08967
hypothetical protein F1915.150 - Arabidopsis thaliana
C1:Species: Arabidopsis thaliana (mouse-ear cress)
C1:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #c
C1:Accession: T08967
R1:Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brand
submitted to the Protein sequence Database, May 1999
A1:Reference number: Z16519
A1:Accession: T08967

A:Molecule type: DNA

A:Residues: 1-334 <BEV>

A:Cross-references: EMBL:AL078470; GSPDB:GN00062; ATSP:FL9B15.150

A:Experimental source: cultivar Columbia; BAC clone FL9B15

C:Genetics:

A:Gene: ATSP:FL9B15.150

A:Map position: 4

C:Superfamily: 3-hydroxyisobutyrate dehydrogenase; 3-hydroxyisobutyrate dehydrogenase homolog <HIB>
F:40-301/Domain: 3-hydroxyisobutyrate dehydrogenase homolog <HIB>

Query Match 46.0%; Score 92; DB 2; Length 334;
Best Local Similarity 45.9%; Pred. No. 0.0001;
Matches 17; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 1 GFLGLGMLGSGIVSNLLKMGHTVTWVNRRTAEKCDLFI 37

DB 41 GWIGTVMGSRMCGHLIKAGYTVTVNRTISKATQLI 77

RESULT 6

G95277

Probable D-threonine [imported] - Sinorhizobium meliloti (strain 1021) megaplasmid pSymA
C:Species: Sinorhizobium meliloti

C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001

C:Accession: G95277

R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001

A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti
A:Reference number: A95262; MUID:21396509; PMID:11481432

A:Accession: G95277

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-293 <KUR>

A:Cross-references: GB:AE006469; PIDN:AAK64785.1; PID:g14523194; GSPDB:GN00165
A:Experimental source: strain 1021, megaplasmid pSymA

R:Galibert, F.; Finan, T.M.; Long, S.R.; Fuhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
Pella, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.

A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104

A:Contents: annotation

C:Genetics:

A:Gene: SMA0237

A:Genome: plasmid

C:Superfamily: 3-hydroxyisobutyrate dehydrogenase; 3-hydroxyisobutyrate dehydrogenase homolog

Query Match 45.0%; Score 90; DB 2; Length 293;

Best Local Similarity 54.5%; Pred. No. 0.00016;

Matches 18; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 2 FLGLGLGMLGSGIVSNLLKMGHTVTWVNRRTAEKCD 34

DB 9 FLGTGLMGAPMARLLGAGFSVTWVNRDAKAE 41

RESULT 7

S14628

phosphoglucuronate dehydrogenase (decarboxylating) (EC 1.1.1.44) - Synecococcus sp.

C:Species: Synecococcus sp.

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: S14628

R:Culler, D.C.; Krogmann, D.W.

submitted to the EMBL Data Library, April 1991

A:Description: Amino acid sequence comparisons of 6-Phosphoglucuronate Dehydrogenase.

A:Reference number: S14628

A:Accession: S14628

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-470 <CUL>

A:Cross-references: EMBL:X58719; NID:947524; PIDN:CAA41555.1; PID:g47525

C:Superfamily: phosphoglucuronate dehydrogenase (decarboxylating); 3-hydroxyisobutyrate de

C:Keywords: oxidoreductase; pentose phosphate pathway

F:6-288/Domain: 3-hydroxyisobutyrate dehydrogenase homolog <HIB>

Query Match 45.0%; Score 90; DB 1; Length 470;

Best Local Similarity 43.2%; Pred. No. 0.00027;

Matches 16; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

QY 1 GFLGLGMLGSGIVSNLLKMGHTVTWVNRRTAEKCDLFI 37

DB 7 GLIGLAVMGENLALNIERNFSLTYNRTAEKTEAFM 43

RESULT 8

H70169

phosphoglucuronate dehydrogenase, decarboxylating (gnd) homolog - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 11-Jun-1999

C:Accession: H70169

R:Praser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White,
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.

Nature 390, 580-586, 1997

A:Authors: Smith, H.O.; Venter, J.C.

A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.

A:Reference number: A70100; MUID:98065943; PMID:9403685

A:Accession: H70169

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-464 <KLE>

A:Cross-references: GB:AE001157; GB:AE000783; NID:92688471; PIDN:AAK66918.1; PID:g2688472

A:Experimental source: strain B31

C:Superfamily: phosphoglucuronate dehydrogenase (decarboxylating); 3-hydroxyisobutyrate de

F:3-280/Domain: 3-hydroxyisobutyrate dehydrogenase homolog <HIB>

Query Match

Best Local Similarity 44.0%; Score 88; DB 2; Length 464;

Matches 16; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

QY 1 GFLGLGMLGSGIVSNLLKMGHTVTWVNRRTAEKCDLFIQ 38

DB 4 GYGLGVMGSLNALNTADNGFNVINRDNEKTEIFVK 41

RESULT 9

G89930

phosphoglucuronate dehydrogenase [imported] - Staphylococcus aureus (strain N315)

C:Species: Staphylococcus aureus

C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001

C:Accession: G89930

R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguchi,
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; H

C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A:Reference number: A89758; MUID:21311952; PMID:11418146

A:Accession: G89930

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-468 <KUR>

A:Cross-references: GB:BA000018; PID:gi3701310; PIDN:BAK42604.1; GSPDB:GN00149

A:Experimental source: strain N315

C:Genetics:

A:Gene: gnd

C:Superfamily: phosphoglucuronate dehydrogenase (decarboxylating); 3-hydroxyisobutyrate de

Query Match

Best Local Similarity 44.0%; Score 88; DB 2; Length 468;

Matches 15; Conservative 12; Mismatches 11; Indels 0; Gaps 0;

QY 1 GFLGLGMLGSGIVSNLLKMGHTVTWVNRRTAEKCDLFIQ 38

DB 1 GFLGLGMLGSGIVSNLLKMGHTVTWVNRRTAEKCDLFIQ 38

RESULT 12
AD1609
6-phosphogluconate dehydrogenase homolog lin1413 [imported] - *Listeria innocua* (strain C)
C;Species: *Listeria innocua* #text change 14-Dec-2001
A;Title: Comparison of whole genome sequences of *Chlamydia pneumoniae* J138.
A;Reference number: A86491; MUID:20330349; PMID:10871362
A;Accession: F86535
A;Status: preliminary

A:Molecule type: DNA
A:Residues: 1-479 <STO>
A:CROSS-references: GB:BA000008; NID:g8978732; PIDN:BAA98568.1; GSPDB:GN00142
A:Experimental source: strain J138
C:Genetics:
A:Gene: gnd
C:Superfamily: phosphogluconate dehydrogenase (decarboxylating); 3-hydroxyisobutyrate de

Query Match 42.5%; Score 85; DB 2; Length 479;
Best Local Similarity 42.1%; Pred. No. 0.0013;
Matches 16; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

QY 1 GFLGLGMSGIVSNLLKMGHTVTVMNRTPAEKCDLFIQ 38
| : | : | : | : | : | : | : | : | : | : | : | : | :
Db 6 GLIGLAVMGKNLVLMIDHGFVSVINRTPEKTRDFLK 43

RESULT 15
C72088
6-phosphogluconate dehydrogenase, decarboxylating CP0398 [imported] - Chlamydia pneumoniae C/Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C:Accession: C72088; H81581
R:Kalan, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.; Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; PMID:99206606; PMID:10192388
A:Accession: C72088
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-479 <ARN>
A:CROSS-references: GB:AE001620; GB:AE001363; NID:g4376631; PIDN:AAD18504.1; PID:g437663
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucl. Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: AB1500; PMID:20150255; PMID:10684935
A:Accession: H81581
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-479 <REA>
A:CROSS-references: GB:AE002201; GB:AE002161; NID:g7189316; PIDN:AAF38243.1; PID:g718932
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucl. Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: AB1500; PMID:20150255; PMID:10684935
A:Accession: H81581
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-479 <REA>
A:CROSS-references: GB:AE002201; GB:AE002161; NID:g7189316; PIDN:AAF38243.1; PID:g718932
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucl. Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: AB1500; PMID:20150255; PMID:10684935
A:Accession: H81581
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-479 <TET>

Query Match 42.5%; Score 85; DB 2; Length 479;
Best Local Similarity 42.1%; Pred. No. 0.0013;
Matches 16; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

QY 1 GFLGLGMSGIVSNLLKMGHTVTVMNRTPAEKCDLFIQ 38
| : | : | : | : | : | : | : | : | : | : | : | : | :
Db 6 GLIGLAVMGKNLVLMIDHGFVSVINRTPEKTRDFLK 43

RESULT 16
AB1714
6-phosphogluconate dehydrogenase, decarboxylating TC0333 [imported] - Chlamydia muridarum C/Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 28-Jul-2000
C:Accession: AB1714
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucl. Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: AB1500; PMID:20150255; PMID:10684935
A:Accession: AB1714
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-479 <TET>

Thu Sep 16 09:36:51 2004

us-10-067-482-2_copy_271_308.rpr

```

A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs4003
C;Superfamily: 3-hydroxyisobutyrate dehydrogenase; 3-hydroxyisobutyrate dehydrogenase h
Query Match 42.2%; Score 84.5; DB 2; Length 299;
Best Local Similarity 32.8%; Pred. No. 0.00092;
Matches 19; Conservative 9; Mismatches 9; Indels 21; Gaps 1;
QY 1 GFLGLGMSGIVSNLLKMGHTVTWNRT-----AEKCDLFI 37
Db 9 GFGLGIMGKPMKXNLLKAGISLVVADRNPEAIADVIAAGTAETASTAKAIAEQCDVII 65

RESULT 19
C85974
probable dehydrogenase yhaE [imported] - Escherichia coli (strain O157:H7, substrain EDL
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: C85974
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouisis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: C85974
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-299 <STO>
A;Cross-references: GB:AE005174; NID:g12517718; PIDN:AAG58255.1; GSPDB:GN00145; UMGp:Z44
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: yhaE
C;Superfamily: 3-hydroxyisobutyrate dehydrogenase; 3-hydroxyisobutyrate dehydrogenase h
Query Match 42.2%; Score 84.5; DB 2; Length 299;
Best Local Similarity 32.8%; Pred. No. 0.00092;
Matches 19; Conservative 9; Mismatches 9; Indels 21; Gaps 1;
QY 1 GFLGLGMSGIVSNLLKMGHTVTWNRT-----AEKCDLFI 37
Db 9 GFGLGIMGKPMKXNLLKAGISLVVADRNPEAIADVIAAGTAETASTAKAIAEQCDVII 66

RESULT 20
C70303
3-hydroxyisobutyrate dehydrogenase - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 11-Jun-1999
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O
V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666; PMID:9537320
A;Accession: C70303
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-288 <AQF>
A;Cross-references: GB:AE000670; NID:g2982779; PIDN:AAC06408.1; PID:g2982783; GB:AE00065
A;Experimental source: strain VF5
C;Genetics:
A;Gene: hbbD
C;Superfamily: 3-hydroxyisobutyrate dehydrogenase; 3-hydroxyisobutyrate dehydrogenase h
F;3-265/Domain: 3-hydroxyisobutyrate dehydrogenase homology <HIB>
Query Match 42.0%; Score 84; DB 2; Length 288;
Best Local Similarity 44.4%; Pred. No. 0.001;
Matches 16; Conservative 5; Mismatches 15; Indels 0; Gaps 0;
QY 1 GFLGLGMSGIVSNLLKMGHTVTWNRTAEKCDLFI 36
Db 4 GFGLGHLGRAIAKRIEQGVBLIWNRTLSKAHEF 39

A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs4003
C;Superfamily: 3-hydroxyisobutyrate dehydrogenase; 3-hydroxyisobutyrate dehydrogenase h
Query Match 42.0%; Score 84; DB 2; Length 295;
Best Local Similarity 33.9%; Pred. No. 0.0011;
Matches 20; Conservative 8; Mismatches 9; Indels 22; Gaps 2;
QY 1 GFLGLGMSGIVSNLLKMGHTVTWNRT-----AEKCDLFI 37
Db 7 GFGLGLMGQGAANILKKGWPLQVNAHNRPAVEMLVAGREAKTPREMAEQCDIV 65

RESULT 22
D96736
probable dehydrogenase F23N20.16 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C;Accession: D96736
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: D96736
A;Status: preliminary
A;Molecule type: DNA
A;Cross-references: GB:AE005173; NID:g6714325; PIDN:AAF26018.1; GSPDB:GN00141
C;Genetics:
A;Gene: F23N20.16
A;Map position: 1
C;Superfamily: 3-hydroxyisobutyrate dehydrogenase; 3-hydroxyisobutyrate dehydrogenase h
Query Match 42.0%; Score 84; DB 2; Length 299;
Best Local Similarity 46.9%; Pred. No. 0.0011;
Matches 15; Conservative 10; Mismatches 7; Indels 0; Gaps 0;
QY 1 GFLGLGMSGIVSNLLKMGHTVTWNRTAEK 32
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2004, 07:31:17 ; Search time 10 Seconds
(without alignments)
197.867 Million cell updates/sec

Title: US-10-067-482-2_COPY_271_308

Perfect score: 200

Sequence: 1 GFLGLMGSGIVSNLLKMGHTVTVMNRTAEKCDLFIQ 38

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	99	49.5	288	1 YKWC_BACSU	O34948 bacillus su
2	90	45.0	470	1 6PGD_SYN7	P21577 synechococ
3	89	44.5	517	1 6PGD_CANAL	O13287 candida alb
4	88	44.0	468	1 6PGD_STAAN	Q99ty2 staphylococ
5	87	43.5	324	1 D3HI_DROME	Q998m5 drosophila
6	87	43.5	468	1 6PGD_STAAM	Q931r3 staphylococ
7	85	42.5	479	1 6PGD_CHLMU	Q9pkx7 chlamydia m
8	85	42.5	479	1 6PGD_CHLPN	Q928i3 chlamydia p
9	84.5	42.2	294	1 GARR_ECOLI	P23523 escherichia
10	84	42.0	488	1 6PGD_TREPA	O83351 treponema p
11	83	41.5	468	1 6PGD_BACSU	P53174 haemophilus
12	81	40.5	467	1 6PGD_BACLI	P12013 bacillus su
13	80	40.0	298	1 YIHU_ECOLI	P52207 bacillus li
14	80	40.0	298	1 YIHU_SALTY	P32142 escherichia
15	80	40.0	479	1 6PGD_TRYBB	Q91780 salmonella
16	80	40.0	484	1 6PGD_HAEIN	P31072 trypanosoma
17	80	40.0	492	1 6PG2_YEAST	P43774 haemophilus
18	79	39.5	468	1 6PG2_BACSU	P53119 saccharomyc
19	79	39.5	480	1 6PGD_CHLTR	P80859 bacillus su
20	79	39.5	489	1 6PG1_YEAST	O84066 chlamydia t
21	77	38.5	484	1 6PGD_ACTAC	P37720 saccharomyc
22	75	37.5	298	1 MMSB_PSEAE	P70718 actinobacil
23	75	37.5	347	1 D3HI_ARATH	P28811 pseudomonas
24	75	37.5	468	1 6PGD_STAEP	Q9suc0 arabidopsis
25	75	37.5	485	1 6PGD_CUNEL	O60037 cunningham
26	74	37.0	468	1 6PG9_ECOLI	P8cp47 staphylococ
27	74	37.0	468	1 6PGD_KLEPN	P37754 escherichia
28	74	37.0	468	1 6PGD_SALTY	P14062 salmonella
29	74	37.0	481	1 6PGD_CERCA	P41570 ceratitis c
30	74	37.0	492	1 6PGD_SCHPO	P78812 schizosacch
31	73	36.5	290	1 Y229_SYNV3	Q55702 synechocyst
32	73	36.5	468	1 6PGD_ECOLI	P00350 escherichia
33	73	36.5	468	1 6PGD_SHIFL	P37756 shigella fl

34	73	36.5	472	1	6PGD_LACLA	Q3chu6 lactococcus
35	73	36.5	472	1	6PGD_LACLC	P96789 lactococcus
36	72	36.0	336	1	D3HI_HUMAN	P31937 homo sapien
37	71	35.5	356	1	GPDA_TREPA	O83973 treponema p
38	69	34.5	335	1	D3HI_RAT	P29266 rattus norv
39	69	34.5	335	1	D3H1_MOUSE	Q991l3 mus musculu
40	68	34.0	332	1	GPDA_CLOAB	Q97i46 clostridium
41	68	34.0	482	1	6PGD_HUMAN	P52209 mus sapien
42	68	34.0	482	1	6PGD_MOUSE	Q9cdc0 mus musculu
43	68	34.0	482	1	6PGD_SHEEP	P00349 ovis aries
44	67	33.5	292	1	GLXR_ECOLI	P77161 escherichia
45	67	33.5	297	1	YFCH_ECOLI	P77775 escherichia
46	66.5	33.2	281	1	HBD_CLODI	P45344 clostridium
47	66	33.0	35	1	D3HI_RABIT	P32185 oryctolagus
48	66	33.0	482	1	6PGD_SINY3	P52208 synechocyst
49	64	32.0	299	1	DH31_CAEEL	Q9xt10 caenorhabdi
50	64	32.0	481	1	6PGD_DROME	P41572 drosophila
51	64	32.0	481	1	6PGD_DROSI	P41573 drosophila
52	63	31.5	334	1	GPDA_CHLPN	Q92751 chlamydia p
53	62	31.0	330	1	GPDA_THETN	Q8r9j3 thermoaer
54	61	30.5	269	1	AROE_AQAE	O67049 aquifex aeo
55	60	30.0	445	1	6PGD_CITAM	P41581 citrobacter
56	60	30.0	445	1	6PGD_CITDI	P41582 citrobacter
57	60	30.0	445	1	6PGD_ESCVU	P41574 escherichia
58	60	30.0	445	1	6PGD_KLEPL	P41575 klebsiella
59	60	30.0	445	1	6PGD_KLETE	P41577 klebsiella
60	60	30.0	473	1	6PGD_BUCAP	Q92hd9 buchnera ap
61	59	29.5	269	1	AROE_ARCFU	O27957 archaeoglob
62	59	29.5	294	1	MMSB_MYCTU	O53814 mycobacteri
63	59	29.5	445	1	6PGD_CITFR	P41583 citrobacter
64	59	29.5	445	1	6PGD_SHIBO	P41578 shigella bo
65	59	29.5	445	1	6PGD_SHIDY	P41579 shigella dy
66	59	29.5	445	1	6PGD_SHISO	P41580 shigella so
67	58.5	29.2	2144	1	GLTI_YEAST	Q12680 saccharomyc
68	58	29.0	314	1	HCDH_HUMAN	Q16836 homo sapien
69	58	29.0	314	1	HCDH_MOUSE	Q61425 mus musculu
70	58	29.0	314	1	HCDH_PIG	P00348 sus scrofa
71	58	29.0	314	1	HCDH_RAT	Q9wvk7 rattus norv
72	58	29.0	334	1	GPDA_CHLCV	Q821m8 chlamydomphi
73	58	29.0	341	1	GPDA_XANAC	Q8pq9 xanthomonas
74	58	29.0	468	1	6PGD_BUCBP	Q89ax5 buchnera ap
75	57.5	28.7	444	1	HEMI_BACAA	Q811c5 bacillus an
76	57.5	28.7	456	1	GUDP_PSEPU	P42205 pseudomonas
77	57	28.5	117	1	Y12K_RHDV	P27412 rabbit hemo
78	57	28.5	117	1	Y12K_RHDV3	P27413 rabbit hemo
79	57	28.5	334	1	GPDA_MYCBO	P59961 mycobacteri
80	57	28.5	340	1	GPDA_ENTFA	P95113 mycobacteri
81	57	28.5	340	1	GPDA_ENTFA	Q834c1 enterococcu
82	57	28.5	468	1	6PGD_BUCAL	P57208 buchnera ap
83	56.5	28.2	444	1	HEMI_BACCR	Q817q8 bacillus ce
84	56.5	28.2	446	1	HEMI_BACME	Q8gcb0 bacillus me
85	56.5	28.2	779	1	SRP_DROME	P52172 drosophila
86	56	28.0	332	1	GPDA_STAAM	Q89u16 staphylococ
87	56	28.0	332	1	GPDA_STAAM	Q8nmw9 staphylococ
88	55	27.5	273	1	CME4_BACSU	P39696 bacillus su
89	55	27.5	296	1	PANE_VIRCH	Q9kpq9 vibrio chol
90	55	27.5	326	1	GPDA_BRAJA	Q89ww0 bradyrhizob
91	55	27.5	334	1	GPDA_CHLMU	Q9pl12 chlamydia m
92	55	27.5	334	1	GPDA_CHLTR	O84719 chlamydia t
93	55	27.5	341	1	GPDA_PSESM	Q883y4 pseudomonas
94	55	27.5	341	1	GPDA_XANCP	Q8pdy0 xanthomonas
95	55	27.5	573	1	ERG1_RAT	P52020 rattus norv
96	54	27.0	223	1	TRKA_STRCO	Q53949 streptomyce
97	54	27.0	331	1	PYRB_PSVT1	Q934t0 psychrobact
98	54	27.0	339	1	GALE_NEIMA	P56997 neisseria m
99	54	27.0	339	1	GALE_NEIMB	P56985 neisseria m
100	54	27.0	339	1	GALE_NEIMC	P56986 neisseria m

ALIGNMENTS

YKWC_BACSU STANDARD; PRT; 288 AA.

ID YKWC_BACSU 15-JUL-1998 (Rel. 36, Created)

AC O34948; 15-JUL-1998 (Rel. 36, Last sequence update)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Hypothetical oxidoreductase ykwc (EC 1.1.1.-).

GN YKWC OR BSU13960.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxID=1423;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RC STRAIN=168;

RA Scanlan E., Devine K.M.;

RA "Sequence of the Bacillus subtilis chromosome from ykua to cse-15.";

RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RC MEDLINE=98044033; PubMed=9384377;

RA Kunz F., Ogasawara N., Mosser I., Albertini A.M., Alloni G.,

RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,

RA Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,

RA Brouillet S., Bruchi C.V., Caldwell B., Capuano V., Carter N.M.,

RA Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,

RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,

RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,

RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,

RA Guisepi G., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,

RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,

RA Joris B., Karamata D., Kasahara Y., Koenigsstein G., Krogh S., Kumano M.,

RA Kobayashi Y., Koetter P., Lardinois S., Lauber J., Lazarevic V.,

RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,

RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,

RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,

RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,

RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,

RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,

RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,

RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,

RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,

RA Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A.,

RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,

RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,

RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,

RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,

RT "The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.";

RT Nature 390:249-256(1997).

CC -!- SIMILARITY: Belongs to the 3-hydroxyisobutyrate dehydrogenase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; AJ222587; CAA10859.1; -.

CC EMBL; Z99111; CAB13269.1; -.

CC PIR; B69870; B69870.

CC Subtilist; BG13328; Ykwc.

CC InterPro; IPR002204; 3hydroxisubut_dh.

CC InterPro; IPR006183; 6PGD.

CC InterPro; IPR006115; 6PGD_NAD.

CC Pfam; PF03446; NAD_binding_2; 1.

CC PRINTS; PR00076; 6PGDHRGNASE.

CC PROSITE; PS00895; 3_HYDROXYISUBUT_DH; 1.

CC EMBL; M55002; AAA27330.1; -.

CC EMBL; X58719; CAA11555.1; -.

CC PIR; S14628; S14628.

CC HSPP; P00349; 2PGD.

CC InterPro; IPR008927; 6GDGH_C-like.

CC InterPro; IPR006183; 6PGD.

CC InterPro; IPR006114; 6PGD_C.

CC InterPro; IPR006113; 6PGD-decarbox.

CC InterPro; IPR006115; 6PGD_NAD.

CC InterPro; IPR006184; 6PGD_BS.

CC Pfam; PF00393; 6PGD; 1.

CC Pfam; PF03446; NAD_binding_2; 1.

CC PRINTS; PR00076; 6PGDHRGNASE.

CC PROSITE; PS00461; 6PGD; 1.

CC Gluconate utilization; Oxidoreductase; Pentose shunt; NADP.

CC EPIVR -> SRSPV (IN REF. 1).

CC R -> A (IN REF. 1).

CC AAERGIPVAFSASLDYFSRSPA ->

CC RQNEFRFRFCFGLRLQPARSLP (IN REF. 2).

CC DYFG -> TTC (IN REF. 1).

CC ERTDRSGS -> KAPIALL (IN REF. 1).

CC CONFLICT 449 452

CC CONFLICT 457 464

KW Hypothetical protein; Oxidoreductase; NAD; Complete proteome.

FT ACT SITE 172 BY SIMILARITY.

SQ SEQUENCE 288 AA; 30711 MW; 976DD9098DB47A30 CRC64;

Query Match 49.5%; Score 99; DB 1; Length 288;

Best Local Similarity 47.4%; Pred. No. 9.1e-06;

Matches 18; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 1 GFGLGLMGSGIVSNLLKMGHTVTVMNRATAEKCDLFQ 38

DB 6 GFGLGLMGSKSMASHILNDGHPVLVYTRYTKKAEISILQ 43

RESULT 2

6PGD_SINP7 STANDARD; PRT; 470 AA.

ID -6PGD_SINP7

AC P21577; 01-MAY-1991 (Rel. 18, Created)

DT 01-MAY-1991 (Rel. 18, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE 6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44).

GN GND.

OS Synecococcus sp. (strain PCC 7942) (Anacystis nidulans R2).

OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.

OX NCBI_TaxID=1140;

RN [1]

RP SEQUENCE FROM N.A., AND SUBSTRATE-BINDING SITE.

RC MEDLINE=90299831; PubMed=2113917;

RA Broedel S.E. Jr., Wolf R.E. Jr.,

RT "Genetic tagging, cloning, and DNA sequence of the Synecococcus sp. strain PCC 7942 gene (gnd) encoding 6-phosphogluconate dehydrogenase.";

RT J. Bacteriol. 172:4023-4031(1990).

RN [2]

RP SEQUENCE FROM N.A.

RA Culler D.C., Krogmann D.W.;

RA Submitted (APR-1991) to the EMBL/GenBank/DBJ databases.

CC -!- CATALYTIC ACTIVITY: 6-phospho-D-gluconate + NADP(+) = D-ribulose 5-phosphate + CO(2) + NADPH.

CC -!- PATHWAY: Hexose monophosphate shunt.

CC -!- SIMILARITY: Belongs to the 6-phosphogluconate dehydrogenase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; M55002; AAA27330.1; -.

CC EMBL; X58719; CAA11555.1; -.

CC PIR; S14628; S14628.

CC HSPP; P00349; 2PGD.

CC InterPro; IPR008927; 6GDGH_C-like.

CC InterPro; IPR006183; 6PGD.

CC InterPro; IPR006114; 6PGD_C.

CC InterPro; IPR006113; 6PGD-decarbox.

CC InterPro; IPR006115; 6PGD_NAD.

CC InterPro; IPR006184; 6PGD_BS.

CC Pfam; PF00393; 6PGD; 1.

CC Pfam; PF03446; NAD_binding_2; 1.

CC PRINTS; PR00076; 6PGDHRGNASE.

CC PROSITE; PS00461; 6PGD; 1.

CC Gluconate utilization; Oxidoreductase; Pentose shunt; NADP.

CC EPIVR -> SRSPV (IN REF. 1).

CC R -> A (IN REF. 1).

CC AAERGIPVAFSASLDYFSRSPA ->

CC RQNEFRFRFCFGLRLQPARSLP (IN REF. 2).

CC DYFG -> TTC (IN REF. 1).

CC ERTDRSGS -> KAPIALL (IN REF. 1).

CC CONFLICT 449 452

CC CONFLICT 457 464

```

FT CONFLICT 468 469 QW -> M (IN REF. 1).
SQ SEQUENCE 470 AA; 50860 MW; 72504AED32594762 CRC64;

Query Match
Best Local Similarity 45.0%; Score 90; DB 1; Length 470;
Matches 16; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

QY 1 GFLGLGLMGSGIVSNLLKMGHTVTVWNTAKCDLFI 37
DB 7 GLIGLAVNGQLNLINMADHGTVTVWNTAKCDLFI 43

RESULT 3
6PGD CANAL STANDARD; PRT; 517 AA.
AC O.3287;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44).
GN DOR14.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
ON NCBI_TaxID=5476;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IFO 1060;
RA Watanabe M., Ishii N., Arisawa M., Aoki Y.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: 6-phospho-D-gluconate + NADP(+) = D-ribulose
CC 5-phosphate + CO(2) + NADPH.
CC -!- PATHWAY: Hexose monophosphate shunt.
CC -!- SIMILARITY: Belongs to the 6-phosphogluconate dehydrogenase
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AB006102; BAA21690.1; -.
CC HSP; P00349; 2PGD.
CC InterPro; IPR008927; 6DGDH_C_like.
CC InterPro; IPR006183; 6PGD.
CC InterPro; IPR006114; 6PGD_C.
CC InterPro; IPR006113; 6PGD_decarbox.
CC InterPro; IPR006115; 6PGD_NAD.
CC Pfam; PF00393; 6PGD; 1.
CC Pfam; PF03446; NAD binding 2; 1.
CC PRINTS; PR00076; 6PGDHRGNASE.
CC TIGRFAMs; TIGR00873; gnd; 1.
CC PROSITE; PS00461; 6PGD; 1.
CC Oxidoreductase; Pentose shunt; NADP.
CC SEQUENCE 517 AA; 56924 MW; 91E3F520FFCABF7A CRC64;

Query Match
Best Local Similarity 44.5%; Score 89; DB 1; Length 517;
Matches 15; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY 1 GFLGLGLMGSGIVSNLLKMGHTVTVWNTAKCDLFI 38
DB 32 GLIGLAVNGQLNLINMADHGTVTVWNTAKCDLFI 69

RESULT 4
6PGD STAAH STANDARD; PRT; 468 AA.
AC Q95TY2;

```

```

DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DE 6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44).
GN GND OR SAI342 OR MW1464.
OS Staphylococcus aureus (strain N315), and
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
ON NCBI_TaxID=158879, 196620;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N315;
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.I., Nagai Y., Llan J., Ito T., Kanamori M.,
RA Matsumaru H., Maruyama A., Murakami H., Hosoyama A., Mizutani-Ui Y.,
RA Takahashi N.K., Sawano T., Inoue R.I., Kaito C., Sekimizu K.,
RA Hiraoka H., Kuhara S., Goto S., Yabuzaki J., Kanehisa M.,
RA Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T., Hattori M.,
RA Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MW2;
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwano M., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";
RL Lancet 359:1819-1827(2002).
CC -!- CATALYTIC ACTIVITY: 6-phospho-D-gluconate + NADP(+) = D-ribulose
CC 5-phosphate + CO(2) + NADPH.
CC -!- PATHWAY: Hexose monophosphate shunt.
CC -!- SIMILARITY: Belongs to the 6-phosphogluconate dehydrogenase
CC family.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AF003134; BAB42604.1; -.
CC EMBL; AF004827; BAB95329.1; -.
CC PIR; G89930; G89930.
CC HSP; P00349; 2PGD.
CC SWISS-2DPAGE; Q99TY2; STAAH.
CC InterPro; IPR008927; 6DGDH_C_like.
CC InterPro; IPR006183; 6PGD.
CC InterPro; IPR006184; 6PGD_BS.
CC InterPro; IPR006114; 6PGD_C.
CC InterPro; IPR006113; 6PGD_decarbox.
CC Pfam; PF00393; 6PGD; 1.
CC Pfam; PF03446; NAD binding 2; 1.
CC PRINTS; PR00076; 6PGDHRGNASE.
CC TIGRFAMs; TIGR00873; gnd; 1.
CC PROSITE; PS00461; 6PGD; 1.
CC Gluconate utilization; Oxidoreductase; Pentose shunt; NADP;
CC Complete proteome.
CC SEQUENCE 468 AA; 51802 MW; 61A5C2CAF3CCD011 CRC64;

Query Match
Best Local Similarity 44.0%; Score 88; DB 1; Length 468;
Matches 15; Conservative 12; Mismatches 11; Indels 0; Gaps 0;

QY 1 GFLGLGLMGSGIVSNLLKMGHTVTVWNTAKCDLFI 38
DB 1 GFLGLGLMGSGIVSNLLKMGHTVTVWNTAKCDLFI 38

```

```

103 6 GVIGLAVNGKNLAWNIESRGYSVFNRSSEKTDLMVE 43
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

```
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekinizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.,
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus."
RT Lancet 357:1225-1240 (2001).
CC -1- CATALYTIC ACTIVITY: 6-phospho-D-gluconate + NADP(+) = D-ribulose
CC 5-phosphate + CO(2) + NADPH.
CC -1- PATHWAY: Hexose monophosphate shunt.
CC -1- SIMILARITY: Belongs to the 6-phosphogluconate dehydrogenase
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF003362; BAB57673.1; -
CC InterPro; IPR008927; 6DGDH_C_like.
CC InterPro; IPR006183; 6PGD.
CC InterPro; IPR006184; 6PGD_BS.
CC InterPro; IPR006114; 6PGD_C.
CC InterPro; IPR006113; 6PGD_decarbox.
CC InterPro; IPR006115; 6PGD_NAD.
CC Pfam; PF00393; 6PGD; 1.
CC Pfam; PF03446; NAD binding 2; 1.
CC PRINTS; PR00076; 6FGDHRGNASE.
CC TIGRFAMS; TIGR00873; gnd; 1.
CC PROSITE; PS00461; 6PGD; 1.
KW Gluconate utilization; Oxidoreductase; Pentose shunt; NADP;
KW Complete proteome.
SQ SEQUENCE 468 AA; 51783 MW; 07205599873133D5 CRC64;

Query Match 43.5%; Score 87; DB 1; Length 468;
Best Local Similarity 39.5%; Pred. No. 0.00057;
Matches 15; Conservative 12; Mismatches 11; Indels 0; Gaps 0;

QY 1 GFLGLGSGIVSNLLKMGHTVTVMNRTAEKCDLFIQ 38
DQ 6 GVIGLAVMGKLNWIESHGVSVFNRSSEKTLNVE 43

RESULT 7
ID 6PGD_CHLMU STANDARD; PRT; 479 AA.
AC Q9PKX7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44).
GN GND OR TC0333.
OS Chlamydia muridarum.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=833560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MoPn / Ni99;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39."
RT Nucleic Acids Res. 28:1397-1406 (2000).
```

```
CC -1- CATALYTIC ACTIVITY: 6-phospho-D-gluconate + NADP(+) = D-ribulose
CC 5-phosphate + CO(2) + NADPH.
CC -1- PATHWAY: Hexose monophosphate shunt.
CC -1- SIMILARITY: Belongs to the 6-phosphogluconate dehydrogenase
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AS002301; AAF39196.1; -
CC PIR; A81714; A81714.
CC HSSP; P00349; 2PGD.
CC TIGR; TC0333; -
CC InterPro; IPR008927; 6DGDH_C_like.
CC InterPro; IPR006183; 6PGD.
CC InterPro; IPR006114; 6PGD_C.
CC InterPro; IPR006113; 6PGD_decarbox.
CC InterPro; IPR006115; 6PGD_NAD.
CC Pfam; PF00393; 6PGD; 1.
CC Pfam; PF03446; NAD binding 2; 1.
CC PRINTS; PR00076; 6FGDHRGNASE.
CC TIGRFAMS; TIGR00873; gnd; 1.
CC PROSITE; PS00461; 6PGD; 1.
KW Gluconate utilization; Oxidoreductase; Pentose shunt; NADP;
KW Complete proteome.
SQ SEQUENCE 479 AA; 52689 MW; 3501DD6DAA0B8F8B CRC64;

Query Match 42.5%; Score 85; DB 1; Length 479;
Best Local Similarity 39.5%; Pred. No. 0.0011;
Matches 15; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

QY 1 GFLGLGSGIVSNLLKMGHTVTVMNRTAEKCDLFIQ 38
DQ 7 GLTGLAVMGKLNVLNMDHGFVSVNRSPEKTEFLK 44

RESULT 8
ID 6PGD_CHLPN STANDARD; PRT; 479 AA.
AC Q9Z8T3; Q9Z8T3.
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44).
GN GND OR CPN0360 OR CP0398 OR CPB0369.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=833558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CWL029;
RX MEDLINE=99206606; PubMed=10192388;
RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis."
RL Nat. Genet. 21:385-389 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39."
RT Nucleic Acids Res. 28:1397-1406 (2000).
```

```

RL Nucleic Acids Res. 28:1397-1406 (2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314 (2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=TW-183;
RA Geng M.M., Schuhmacher A., Muehldorfer I., Bensch K.W., Schaefer K.P.,
RA Schneider S., Pohl T., Essig A., Marre R., Melchers K.;
RT "The genome sequence of Chlamydia pneumoniae TW183 and comparison with
RT other Chlamydia strains based on whole genome sequence analysis.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC !- CATALYTIC ACTIVITY: 6-phospho-D-gluconate + NADP(+) = D-ribulose
CC 5-phosphate + CO(2) + NADPH.
CC !- PATHWAY: Hexose monophosphate shunt.
CC !- SIMILARITY: Belongs to the 6-phosphogluconate dehydrogenase
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AE001620; AAD18504.1; -
CC EMBL; AE002201; AAF38243.1; -
CC EMBL; AP002546; BAA98568.1; -
CC EMBL; AE017158; AAP98300.1; -
CC PIR; C72088; C72088
CC PIR; F86535; F86535.
CC HSSP; P00349; 2PGD.
CC TIGR; CP0398; -
CC InterPro; IPR008927; 6GDH_C_like.
CC InterPro; IPR006183; 6PGD.
CC InterPro; IPR006114; 6PGD_C.
CC InterPro; IPR006113; 6PGD_decarbox.
CC InterPro; IPR006115; 6PGD_NAD.
CC Pfam; PF00393; 6PGD; 1.
CC Pfam; PF03446; NAD_binding_2; 1.
CC PRINTS; PR00076; 6PGDHDRGNASE.
CC TIGRFAMs; TIGR00873; gnd; 1.
CC PROSITE; PS00461; 6PGD; 1.
CC Gluconate utilization; Oxidoreductase; Pentose shunt; NADP;
KW Complete proteome.
SQ SEQUENCE 479 AA; 52938 MW; 0FD301D3378E11FA CRC64;

Query Match 42.5%; Score 85; DB 1; Length 479;
Best Local Similarity 42.1%; Pred. No. 0.0011;
Matches 16; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

QY 1 GFLGLMGSGIVGNLLKMGHTVTVWNRATFKCDLFIQ 38
DB 6 GLIGLAVGNKLVLMIDHGFVSVYNTPEKTRDFLK 43

RESULT 9
GARR_ECOLI
ID GARR_ECOLI STANDARD; PRT; 294 AA.
AC 23523;
DT 01-NOV-1991 (Rel. 20, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 2-hydroxy-3-oxopropionate reductase (EC 1.1.1.60) (Tartronate
DE semialdehyde reductase) (TSAR).

```

```

GN GARR OR B3125 OR C3880.
OS Escherichia coli, and
OS Escherichia coli O6
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 217992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=91154140; PubMed=1705543;
RA Komine Y., Inokuchi H.;
RT "Precise mapping of the rnpB gene encoding the RNA component of RNase
RT P in Escherichia coli K-12.";
RL J. Bacteriol. 173:1813-1816 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97426617; PubMed=9279503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474 (1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=98447507; PubMed=9772162;
RA Hubbard B.K., Koch M., Palmer D.R., Babbitt P.C., Gerlt J.A.;
RT "Evolution of enzymatic activities in the enolase superfamily:
RT characterization of the (D)-glucarate/galactarate catabolic pathway
RT in Escherichia coli.";
RL Biochemistry 37:14369-14375 (1998).
RN [5]
RP GENE NAME.
RX MEDLINE=20225875; PubMed=10762278;
RA Monterrubio R., Baldoma L., Obradors N., Aguilar J., Badia J.;
RT "A common regulator for the operons encoding the enzymes involved in
RT D-galactarate, D-glucarate, and D-glycerate utilization in
RT Escherichia coli.";
RL J. Bacteriol. 182:2672-2674 (2000).
CC !- CATALYTIC ACTIVITY: (R)-glycerate + NAD(P)(+) = 2-hydroxy-3-
CC oxopropanoate + NAD(P)H.
CC !- PATHWAY: D-galactarate metabolism; third step.
CC !- SIMILARITY: Belongs to the 3-hydroxyisobutyrate dehydrogenase
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; D90212; BAA14238.1; ALT_INIT.
CC EMBL; U18997; AAA57928.1; ALT_INIT.
CC EMBL; AE000394; AAC76159.1; ALT_INIT.
CC EMBL; AE016767; AAN82321.1; ALT_INIT.
CC EcoGene; EG11176; gar.
CC InterPro; IPR002204; 3hydroxisobut_dh.
CC InterPro; IPR006183; 6PGD.
CC InterPro; IPR006115; 6PGD_NAD.

```

KW	Gluconate utilization; Oxidoreductase; Pentose shunt; NADP;
KX	Complete proteome.
QW	SEQUENCE 488 AA; 52767 MW; FB69CCCA98DE6B5 CRC64;
QY	Query Match 42.08; Score 84; DB 1; Length 488;
DB	Best Local Similarity 43.2%; Pred. NO. 0.0015;
	Matches 16; Conservative 8; Mismatches 13; Indels 0; Gaps
1	GFLGLGLMGSGIVSNLLKMGHTVTVMNRATAEKCDLFI 37
6	GFGLVAGNGLVNIERNFGSVAVNRITVTVDREL 42
RESULT 11	
6PGD	BACSU
AD	6PGD BACSU STANDARD; PRT; 468 AA.
IC	P12013;
DT	01-OCT-1989 (Rel. 12, Created)
DT	01-OCT-1989 (Rel. 12, Last sequence update)
DT	10-OCT-2003 (Rel. 42, Last annotation update)
DE	6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44).
GN	GNTZ OR BSU40080.
OS	Bacillus subtilis.
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX	NCBI_TaxID=1423;
RP	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=87008613; PubMed=3020045;
RX	Fujita Y., Fujita T., Miwa Y., Nihashi J., Aratani Y.;
RT	"Organization and transcription of the gluconate operon, gnt, of
RT	Bacillus subtilis.;"
RL	J. Biol. Chem. 261:13744-13753(1986).
RL	[2]
RP	SEQUENCE FROM N.A.
RC	STRAIN=168 / BGSC141;
RC	MEDLINE=96093926; PubMed=7584049;
RX	Yoshida K.-I., Seki S., Fujimura M., Miwa Y., Fujita Y.;
RT	"Cloning and sequencing of a 36-kb region of the Bacillus subtilis
RT	genome between the gnt and iol operons.;"
RL	DNA Res. 2:61-69(1995).
RL	[3]
RP	SEQUENCE FROM N.A.
RP	STRAIN=168;
RC	MEDLINE=98044033; PubMed=9384377;
RA	Kunst F., Ogasawara N., Moser I., Albertini A.M., Alloni G.,
RA	Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA	Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA	Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA	Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,
RA	Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA	Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA	Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA	Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA	Guiseppi G., Guy B.-J., Hage K., Halech J., Harwood C.R., Henaut A.,
RA	Hilbert H., Holsappel S., Hosono S., Kullio M.F., Itaya M., Jones L.,
RA	Joris B., Karamata D., Kasahara Y., Kluwe-Blanchard M., Klein C.,
RA	Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA	Kurita K., Lapidus A., Lardinols S., Lauber J., Lazarevic V.,
RA	Lee S.M., Levine A., Liu H., Masuda S., Maul C., Medigic C.,
RA	Medina N., Meliardo R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
RA	Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA	Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA	Presecan E., Puig C., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA	Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA	Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA	Seiguchi J., Sekowska A., Seror S.J., Serro P., Shin B.S., Soldo B.,
RA	Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA	Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA	Toato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA	Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA	Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA	Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RT	"The complete genome sequence of the Gram-positive bacterium Bacillus


```
RT subtilis.";
RL Nature 390:249-256 (1997).
RN [4]
RP SEQUENCE OF 460-468 FROM N.A.
RC STRAIN=168;
RA Kasahara Y., Nakai S., Yoshikawa H., Ogasawara N.;
RT "36kb sequence between gntZ and tny of B. subtilis genome.";
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RN [5]
RP PROBABLE FUNCTION.
RX MEDLINE=92065803; PubMed=1659648;
RA Reizer A., Deutscher J., Saier M.H. Jr., Reizer J.;
RT "Analysis of the Gluconate (gnt) operon of Bacillus subtilis.";
RL Mol. Microbiol. 5:1081-1089 (1991).
CC -|- CATALYTIC ACTIVITY: 6-phospho-D-gluconate + NADP(+) = D-ribulose
CC 5-phosphate + CO(2) + NADPH.
CC -|- PATHWAY: Hexose monophosphate shunt.
CC -|- SIMILARITY: Belongs to the 6-phosphogluconate dehydrogenase
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; J02584; AAA56927.1; -.
CC DR EMBL; AB005554; BAA21576.1; -.
CC DR EMBL; Z91124; CAB16045.1; -.
CC DR EMBL; D78193; BAA11267.1; -.
CC DR PIR; D26190; D26190.
CC DR HSP; P00349; 2PGD.
CC DR Subtilisin; EG10651; gntZ.
CC DR InterPro; IPR008927; 6GDGH_C-like.
CC DR InterPro; IPR006183; 6PGD.
CC DR InterPro; IPR006114; 6PGD C.
CC DR InterPro; IPR006113; 6PGD decarbox.
CC DR InterPro; IPR006115; 6PGD_NAD.
CC DR Pfam; PF03446; NAD_binding_2; 1.
CC DR PRINTS; PR00076; 6EGDHRGNASE.
CC DR TIGRFAMs; TIGR00873; gnd; 1.
CC DR PROSITE; PS00461; 6PGD; 1.
CC DR Gluconate utilization; Oxidoreductase; Pentose shunt; NADP;
CC Complete proteome.
CC SEQUENCE 468 AA; 51983 MW; 56D88BEBE553856 CRC64;

Query Match 41.5%; Score 83; DB 1; Length 468;
Best Local Similarity 44.7%; Pred. No. 0.0019;
Matches 17; Conservative 5; Mismatches 16; Indels 0; Gaps 0;

Oy 1 GFLGLGMSGIVSNLLKMGHTVTVWNRATKCDLFIQ 38
Db 6 GVIGLGVGNSIALNMKGENVAVNYTRDLTDLIQ 43

RESULT 12
6PGD_BACLI STANDARD; PRT; 467 AA.
AC P52207;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44).
GN GNTZ.
OS Bacillus licheniformis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1402;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RC STRAIN=BGSC5A2;
RX MEDLINE=96051988; PubMed=8535972;
RA Yoshida K., Seki S., Fujita Y.;
RT "Nucleotide sequence and features of the Bacillus licheniformis gnt
RT operon.";
RL DNA Res. 1:157-162 (1994).
CC -|- CATALYTIC ACTIVITY: 6-phospho-D-gluconate + NADP(+) = D-ribulose
CC 5-phosphate + CO(2) + NADPH.
CC -|- PATHWAY: Hexose monophosphate shunt.
CC -|- SIMILARITY: Belongs to the 6-phosphogluconate dehydrogenase
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; D31631; BAA06504.1; -.
CC DR PIR; JC2306; JC2306.
CC DR HSP; P00349; 2PGD.
CC DR InterPro; IPR008927; 6GDGH_C-like.
CC DR InterPro; IPR006183; 6PGD.
CC DR InterPro; IPR006114; 6PGD C.
CC DR InterPro; IPR006113; 6PGD decarbox.
CC DR InterPro; IPR006115; 6PGD_NAD.
CC DR InterPro; IPR006184; 6PGD_BS.
CC DR Pfam; PF03446; NAD_binding_2; 1.
CC DR PRINTS; PR00076; 6EGDHRGNASE.
CC DR TIGRFAMs; TIGR00873; gnd; 1.
CC DR PROSITE; PS00461; 6PGD; 1.
CC DR Gluconate utilization; Oxidoreductase; Pentose shunt; NADP.
CC SEQUENCE 467 AA; 51642 MW; 3F14CF0320D8E52 CRC64;

Query Match 40.5%; Score 81; DB 1; Length 467;
Best Local Similarity 42.1%; Pred. No. 0.0035;
Matches 16; Conservative 6; Mismatches 16; Indels 0; Gaps 0;

Oy 1 GFLGLGMSGIVSNLLKMGHTVTVWNRATKCDLFIQ 38
Db 6 GVIGLGVGNSIALNMKGENVAVNYTRDLTDLIQ 43

RESULT 13
YIHU_ECOLI STANDARD; PRT; 298 AA.
ID YIHU_ECOLI
AC P32142;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical oxidoreductase yihU (EC 1.1.1.-).
GN YIHU OR B3882 OR SF3954 OR S3792.
OS Escherichia coli, and
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 623;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12 / MG1655;
RX MEDLINE=93347969; PubMed=8346018;
RA Plunkett G. III, Burland V.D., Daniels D.L., Blattner F.R.;
RT "Analysis of the Escherichia coli genome. III. DNA sequence of the
RT region from 87.2 to 89.2 minutes.";
RL Nucleic Acids Res. 21:3391-3398 (1993).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
```


	Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.
RA	Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA	Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA	Yu J.;
RT	"Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT	through comparison with genomes of Escherichia coli K12 and O157";
RL	Nucleic Acids Res. 30:4432-4441(2002).
RN	[3]
CC	SEQUENCE FROM N.A.
RP	SPECIES=S.flexneri; STRAIN=2457T / ATCC 700930 / Serotype 2a;
RC	MEDLINE=22590274; PubMed=12704152;
RX	Weil J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA	Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
MA	Bau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA	Schwartz D.C., Blattner F.R.;
RT	"Complete genome sequence and comparative genomics of Shigella
RT	flexneri serotype 2a strain 2457T.";
RL	Infect. Immun. 71:2775-2786(2003).
CC	-! SIMILARITY: Belongs to the 3-hydroxyisobutyrate dehydrogenase family.

CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/about or send an email to licenses@isb-sib.ch)
EMBL	L19201; AAC03015.1; --
DR	EMBL; AB000464; AAID3444.1; --
DR	EMBL; AE015402; AAN45389.1; --
DR	PIR; AE016990; AAP18811.1; --
DR	EPR; SA0826; SA0826.
DR	EcoGene; IGI1847; yihU.
DR	InterPro; IPR002204; 3hydroxisubut_dh.
DR	InterPro; IPR006115; 6PGD_NAD.
DR	Pfam; PF03446; NAD binding_2; 1.
DR	PROSITE; PS00895; 3_HYDROXYISOBUT_DH; 1.
KW	Hypothetical protein; Oxidoreductase; NAD; Complete proteome.
FT	ACT_SITE 171..171 BY SIMILARITY.
SEQ	FEATURES 171..171
SC	SEQUENCE 298 AA; 31158 MW; 74FBC8C9FA7981C CRC64;
Query Match	40.0%; Score 80; DB 1; Length 298;
Best Local Similarity	53.3%; Pred.No. 0.0031;
Matches 16;	Conservative 6; Mismatches 8; Indels 0; Gaps
QY	2 FTGLMGSGIVSNLLKXGHTVTWNRTAE 31 - : : : :
Dd	6 FTGLGGSPMASNLQQCHQLRFVDVNAE 35 - : : : :
RESULT 14	
YIHU_SALTY	
ID	YIHU_SALTY STANDARD; PRT; 298 AA.
AC	O9L7SO;
DT	16-OCT-2001 (Rel. 40, Created)
DD	16-OCT-2001 (Rel. 40, Last sequence update)
DE	28-FEB-2003 (Rel. 41, Last annotation update)
DE	Hypochemical oxidoreductase yihu (EC 1.1.-.-.).
GK	yihu OR STM4023.
GN	yihu
OS	Salmonella typhimurium.
OC	Bacteria; Proteobactera; Gammaproteobacteria; Enterobacteriales;
OX	Enterobacteriaceae; Salmonella.
NCBI_TaxID=602;	
(1)	
RN	SEQUENCE FROM N.A.
RC	STRAIN=LT2;
RP	
RA	Krogan N.J., Zhang R., Neuhaud J., Kelln R.A.;
RT	"Utilization of dihydrorotate as sole pyrimidine source by Salmonella typhimurium."
RL	Submitted (JAN-2000) to the EMBL/GenBank/DBAJ databases.
RN	[2]

SEQUENCE FROM N.A.

RP STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856 (2001).
CC -!- SIMILARITY: Belongs to the 3-hydroxyisobutyrate dehydrogenase
CC family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licenses@isb-sib.ch).

DR ENBL; AF220438; AAF27921.1; -;
DR ENBL; AB008887; AAL22862.1; -;
DR StyGene; SG?????; YiHo.
DR InterPro; IPR002204; 3hydroxisobut_dh.
DR InterPro; IPR006115; 6PGD NAD.
DR InterPro; IPR000205; NAD BS.
DR Pfam; PF03446; NAD binding 2, 1.
DR PROSITE; PS00895; 3-HYDROXYISOBUT_DH; 1.
DR Hypothetical protein; Oxidoreductase; NAD; Complete proteome.
KW ACT SITE 171 BY SIMILARITY.
FT ACT SITE 171
SQ SEQUENCE 298 AA; 31280 MW; 7ED9FA174249240C CRC64;

Query Match 40.0%; Score 80; DB 1; Length 298;
Best Local Similarity 57.7%; Pred. No. 0.0031;
Matches 15; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 2 FLGLGLMGSGIVSNLLKMGHTVTWN 27
|::|||::|||::|||::|||::|||:
DB 6 FLGLGQGSPMASNLLKQGHQLSVFD 31
|::|||::|||::|||::|||::|||:

RESULT 15

6PGD_TRYBB STANDARD; PRT; 479 AA.

ID _PGD_TRYBB
AC P31072;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44).
GN GND.
OS Trypanosoma brucei brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=427.
RX MEDLINE=93149212; PubMed=8426618;
RA Barrett M.P., le Page R.W.F.;
RT "A 6-phosphogluconate dehydrogenase gene from Trypanosoma brucei."
RI Mol. Biochem. Parasitol. 57:89-100(1993).
RL [2]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
RX MEDLINE=98411456; PubMed=9737929;
RA Phillips C., Dohnalek J., Gover S., Barrett M.P., Adams M.J.;
RT "A 2.8-A resolution structure of 6-phosphogluconate dehydrogenase
RT from the protozoan parasite Trypanosoma brucei: comparison with the
RT sheep enzyme accounts for differences in activity with coenzyme and
RT substrate analogues.";
RJ J. Mol. Biol. 282:667-681(1998).
CC -!- CATALYTIC ACTIVITY: 6-phospho-D-gluconate + NADP(+) = D-ribulose
CC 5-phosphate + CO(2) + NADPH.

-1- PATHWAY: Hexose monophosphate shunt.
CC CC -1- SUBUNIT: Homodimer.
CC CC -1- SIMILARITY: Belongs to the 6-phosphogluconate dehydrogenase family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>) or send an email to license@isb-sib.ch.

EMBL; X65623; CAA48577.1; --
DR PIR; A48565; A48565.
DR PDB; 1PGJ; 11-NOV-98.
DR InterPro; IPR008927; 6DGDH_C-like.
DR InterPro; IPR006183; 6PGD.
DR InterPro; IPR006114; 6PGD_C.
DR InterPro; IPR006113; 6PGD_decarbox.
DR InterPro; IPR006115; 6PGD_NAD.
DR InterPro; IPR006184; 6PGdom_BS.
DR Pfam; PF03446; NAD binding 2; 1.
DR PRINTS; PR00076; 6PGDHGRNASE.
DR TIGRfams; TIGR00873; gnd; 1.
DR PROSITE; PS00461; 6PGD_1.
KW Gluconate utilization; Oxidoreductase; Pentose shunt; NADP;
KM 3D-structure.

STRAND	3	3
FT STRAND	5	8
FT HELIX	12	23
FT TURN	24	25
FT STRAND	28	31
FT HELIX	35	44
FT TURN	45	47
FT TURN	49	50
FT HELIX	51	53
FT STRAND	54	56
FT HELIX	60	66
FT STRAND	67	67
FT STRAND	72	75
FT HELIX	81	93
FT TURN	96	97
FT STRAND	99	102
FT HELIX	108	119
FT TURN	120	122
FT STRAND	124	131
FT HELIX	133	139
FT STRAND	142	147
FT HELIX	149	162
FT STRAND	165	165
FT TURN	167	168
FT STRAND	171	171
FT TURN	179	180
FT HELIX	181	209
FT TURN	210	211
FT HELIX	214	226
FT TURN	227	227
FT TURN	229	230
FT STRAND	232	232
FT HELIX	233	243
FT STRAND	245	245
FT TURN	247	248
FT STRAND	251	251
FT HELIX	252	255
FT STRAND	256	256
FT HELIX	265	276
FT TURN	277	277
FT HELIX	281	294
FT TURN	295	295
FT HELIX	296	305
FT TURN	307	310

```
DR PIR; C64077; C64077.
DR HSSP; P00349; 2PGD.
DR TIGR; H10553; -.
DR InterPro; IPR008927; 6DGDH_C_like.
DR InterPro; IPR006183; 6PGD.
DR InterPro; IPR006114; 6PGD C.
DR InterPro; IPR006113; 6PGD decarbox.
DR InterPro; IPR006115; 6PGD NAD.
DR Pfam; PF00393; 6PGD; 1.
DR PRINTS; PR00076; NAD binding 2; 1.
DR InterPro; IPR006184; 6PGDm_BS.
DR Pfam; PF00393; 6PGD; 1.
DR PRINTS; PR00076; NAD binding 2; 1.
DR PROSITE; PS00461; 6PGD; 1.
DR TIGRFAMs; TIGR00873; gnd; 1.
DR PROSITE; PS00461; 6PGD; 1.
DR Oxidoreductase; Pentose shunt; NADP.
KW Complete proteome.
SQ SEQUENCE 484 AA; 53140 MW; 8381EEB3C704C5DA CRC64;

Query Match 40.0%; Score 80; DB 1; Length 484;
Best Local Similarity 39.5%; Pred. No. 0.0049;
Matches 15; Conservative 7; Mismatches 16; Indels 0; Gaps 0;

QY 1 GFLGLGMSGIVSNLLKMGHTVTVMNRTAEKCDLFIQ 38
DB 8 GVIGLAVMGQNLILNMNDHGFVKVAYNRTTSKYVDFLQ 45

RESULT 17
6PG2_YEAST
ID 6PG2_YEAST STANDARD; PRT; 492 AA.
AC P53319;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 6-phosphogluconate dehydrogenase, decarboxylating 2 (EC 1.1.1.44).
GN GND2 OR YGR256W OR G9170.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c;
RX MEDLINE=97279233; PubMed=9133741;
RA Mazzoni C., Ruzzi M., Rinaldi T., Solinas F., Montebove F.,
RA Frontali L.;
RT "Sequence analysis of a 10.5 kb DNA fragment from the yeast
RT chromosome VII reveals the presence of three new open reading frames
RT and of a tRNAThr gene.";
RL Yeast 13:369-372(1997).
CC -!- CATALYTIC ACTIVITY: 6-phospho-D-gluconate + NADP(+) = D-ribulose
CC 5-phosphate + CO(2) + NADPH.
CC -!- PATHWAY: Hexose monophosphate shunt.
CC -!- SIMILARITY: Belongs to the 6-phosphogluconate dehydrogenase
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X99228; CAA57612.1; -.
CC DR EMBL; Z73041; CAA97285.1; -.
CC PIR; S64588; S64588.
CC DR HSSP; P00349; 2PGD.
CC GerMOnline; 141568; -.
CC DR SGD; S0003488; GND2.
CC DR GO; GO:0004616; F:6-phosphogluconate dehydrogenase (decarboxyla. .; IMP.
CC DR GO; GO:0006006; P:glucose metabolism; IGI.
CC InterPro; IPR008927; 6DGDH_C_like.
```

```
DR InterPro; IPR006183; 6PGD.
DR InterPro; IPR006114; 6PGD C.
DR InterPro; IPR006113; 6PGD decarbox.
DR InterPro; IPR006115; 6PGD_NAD.
DR Pfam; PF00393; 6PGD; 1.
DR PRINTS; PR00076; NAD binding 2; 1.
DR TIGRFAMs; TIGR00873; gnd; 1.
DR PROSITE; PS00461; 6PGD; 1.
DR Oxidoreductase; Pentose shunt; NADP.
KW Complete proteome.
SQ SEQUENCE 492 AA; 53922 MW; 3D75D53563987735 CRC64;

Query Match 40.0%; Score 80; DB 1; Length 492;
Best Local Similarity 40.5%; Pred. No. 0.005;
Matches 15; Conservative 6; Mismatches 16; Indels 0; Gaps 0;

QY 1 GFLGLGMSGIVSNLLKMGHTVTVMNRTAEKCDLFI 37
DB 9 GLVGLAVMGQNLILNAADHGFTVAYNRTQSKVDRL 45

RESULT 18
6PG2_BACSU
ID 6PG2_BACSU STANDARD; PRT; 468 AA.
AC P80859; P54546;
DT 01-OCT-1996 (Rel. 34, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 6-phosphogluconate dehydrogenase, decarboxylating II (EC 1.1.1.44)
DE (GNT2II).
GN YQJ1 OR BSU23860.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / JH642;
RX MEDLINE=97124195; PubMed=8969508;
RA Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M.,
RA Kobayashi Y.;
RT "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of
RT the Bacillus subtilis genome containing the skin element and many
RT sporulation genes.";
RL Microbiology 142:3103-3111(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Barbero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Bouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Konigstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Solido B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
```


SEQUENCE FROM N.A.
RA Desouza M., Lobo Z., Maitra P.K.;
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
[2]
SEQUENCE FROM N.A.
RA STRAIN=S288c / AB972;
RC MEDLINE=94378003; PubMed=8091229;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
RA Du Z., Favello A., Fulton L., Gattung S., Geisel C., Kirsten J.,
RA Kucaba T., Hillier L.W., Jier M., Johnston L.L., Langston Y.,
RA Latrelle P., Louis E.J., Macri C., Mardis E., Meneses S., Mouser L.,
RA Nhan M., Rifkin L., Riles L., St Peter H., Trevaskis E., Vaughan K.,
RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
RA Vauadin M.;
RT "Complete nucleotide sequence of *Saccharomyces cerevisiae* chromosome
VIII."; Science 265:2077-2082(1994).
[3]
SEQUENCE OF 41-47 AND 120-131.
RA STRAIN=S288c;
RC MEDLINE=95203288; PubMed=7895733;
RA Garrels J.I., Fitcher B., Kobayashi R., Latter G.I., Schwender B.,
RA Volpe T., Warner J.R., McLaughlin C.S.;
RT "Protein identifications for a *Saccharomyces cerevisiae* protein
database."; Electrophoresis 15:1466-1486(1994).
RA Electrophoresis 15:1466-1486(1994).
CC -|- CATALYTIC ACTIVITY: 6-phospho-D-gluconate + NADP(+) = D-ribulose
5-phosphate + CO(2) + NADPH.
CC -|- PATHWAY: Hexose monophosphate shunt.
CC -|- SIMILARITY: Belongs to the 6-phosphogluconate dehydrogenase
family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; Z46631; CAA86600.1; -;
DR EMBL; U00028; AAB68452.1; -;
DR EMBL; U17155; AAA53637.1; -;
DR PIR; S46671; S46671.
DR HSSP; P00349; 2PGD.
DR GermOnline; 139501; -;
DR SGD; S0001226; GND1.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0004616; F:6-phosphogluconate dehydrogenase (decarboxyla. .; IMP.
DR GO; GO:0006006; P:glucose metabolism; IGI.
DR InterPro; IPR008927; 6GDGH_C_like.
DR InterPro; IPR006183; 6PGD.
DR InterPro; IPR006114; 6PGD_C.
DR InterPro; IPR006113; 6PGD decarbox.
DR InterPro; IPR006115; 6PGD_NAD.
DR InterPro; IPR006184; 6PGdom_BS.
DR Pfam; PF00393; 6PGD; 1.
DR Pfam; PF03446; NAD binding 2; 1.
DR PRINTS; PR00076; 6PGDHDHGNASE.
DR TIGRFAMS; TIGR00873; gnd; 1.
DR PROSITE; PS00461; 6PGD; 1.
KW Oxidoreductase; Pentose shunt; NADP.
SQ SEQUENCE 489 AA; 53543 MW; 55CAE5DACDC6A00B CRC64;

Query Match 39.5%; Score 79; DB 1; Length 489;
Best Local Similarity 40.5%; Pred. No. 0.0067;
Matches 15; Conservative 6; Mismatches 16; Indels 0; Gaps 0;

QY 1 GFTGLGIMGSGIVSNLLKMGHTVTVMNRTAEKCDLFI 37
Db 6 GLTGLVGMGNLLINAADHGFTVTCVNNRTOSKVDHFL 42

```

RESULT 21
ID   _6PGD_ACTAC STANDARD; PRT; 484 AA.
AC   P70718;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44).
GN   GND.
OS   Actinobacillus actinomycetemcomitans (Haemophilus
OC   actinomycetemcomitans).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Actinobacillus.
OX   NCBI_TaxID=714;
FN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Y4;
RX   MEDLINE=97148607; PubMed=9020051;
RA   Yoshida Y., Nakano Y., Yamashita Y., Koga T.;
RT   "The gnd gene encoding a novel 6-phosphogluconate dehydrogenase and
RT   its adjacent region of Actinobacillus actinomycetemcomitans
RT   chromosomal DNA."
RL   Biochem. Biophys. Res. Commun. 230:220-225(1997).
CC   -! CATALYTIC ACTIVITY: 6-phospho-D-gluconate + NADP(+)= D-ribulose
CC   5-phosphate + CO(2)+NADPH.
CC   -! PATHWAY: Hexose monophosphate shunt.
CC   -! SIMILARITY: Belongs to the 6-phosphogluconate dehydrogenase
CC   family.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC   European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@lsb-sib.ch).
CC   -----
EMBL; D88189; BAA13558.1; -.
DR   PIR; JC5282; JCS282.
DR   HSPG; P00349; 2PGD.
DR   InterPro; IPR008927; 6GDH_C_like.
DR   InterPro; IPR006183; 6PGD.
DR   InterPro; IPR006114; 6PGD_C.
DR   InterPro; IPR006113; 6PGD_decarb.
DR   InterPro; IPR006115; 6PGD_NAD.
DR   InterPro; IPR006184; 6PGDom_BS.
DR   Pfam; PF00393; 6PGD; 1.
DR   Pfam; PF03446; NAD_binding_2; 1.
DR   PRINTS; PR00076; 6PGDHDRGNASE.
DR   TIGRFAMs; TIGR00873; gnd; 1.
DR   PROSITE; PS00461; 6PGD; 1.
KW   Gluconate utilization; Oxidoreductase; Pentose shunt; NADP.
SQ   SEQUENCE 484 AA; 53288 MW; F7BD0B2EA3BF624D CRC64;

Query Match          38.5%; Score 77; DB 1; Length 484;
Best Local Similarity 36.8%; Pred.No. 0.012;
Matches 14; Conservative      8; Mismatches 16; Indels    0; Gaps    0;

QY   1 GFGLGMGSIGIVSNLLKMGHTVTWNRTAEKCDFIQ 38
     |||::||::||:|||||::|||::|
DB   8 GVIGLAMVGQNLIINMNDHGFKVAVYNRTTSKVDFLE 45
     |::|||::||:|||||::|||::|

RESULT 22
MMGB_PSEAE
ID   MMBS_PSEAE STANDARD; PRT; 298 AA.
AC   P28811;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   3-hydroxyisobutyrate dehydrogenase (EC 1.1.1.31) (HIBADH) .
GN   MMBS OR PA3569.
OS   Pseudomonas aeruginosa.
```

```

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PAO1;
RX MEDLINE=92317087; PubMed=1339433;
RA Steele M.I., Lorenz D., Hatter K., Park A., Sokatch J.R.;
RT "Characterizing the mmsAB operon of Pseudomonas aeruginosa PAO
RT encoding methylmalonate-semialdehyde dehydrogenase and 3-
RT hydroxyisobutyrate dehydrogenase.";
RL J. Biol. Chem. 267:13585-13592(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PAO1;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
CC -!- CATALYTIC ACTIVITY: 3-hydroxy-2-methylpropanoate + NAD(+) = 2-
CC methyl-3-oxopropanoate + NADH.
CC -!- PATHWAY: Distal valine metabolic pathway.
CC -!- INDUCTION: By valine.
CC -!- SIMILARITY: Belongs to the 3-hydroxyisobutyrate dehydrogenase
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M84911; AAA25892.1; -.
DR EMBL; AE004778; AAG06957.1; -.
DR PIR; C42902; C42902.
DR InterPro; IPR002204; 3hydroxisobut_dh.
DR InterPro; IPR006115; 6FGD_NAD.
DR Pfam; PF03446; NAD_binding_2; 1.
DR PROSITE; PS00895; 3_HYDROXYISOBUT_DH; 1.
KW Oxidoreductase; NAD; Complete proteome.
FT NP_BIND 3 31 NAD (BY SIMILARITY).
FT ACT_SITE 171 171 BY SIMILARITY.
SQ SEQUENCE 298 AA; 30486 MW; 0C4D7B5A7C870730 CRC64;

Query Match 37.5%; Score 75; DB 1; Length 298;
Best Local Similarity 57.7%; Pred. NO. 0.014;
Matches 15; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 2 FLGLGMSGIVSNLLKMGHTVTVWN 27
Db 6 FLGLGNMGPGMAANLLKAGHRVNVFD 31

RESULT 23
D3H1_ARATH
ID D3H1_ARATH STANDARD; PRT; 347 AA.
AC Q9SUC0; Q8LUC25;
DT 16-OCT-2001 (Rel. 40, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Probable 3-hydroxyisobutyrate dehydrogenase, mitochondrial precursor
DE (EC 1.1.1.31) (HIBADB).
GN AT4G20930 OR T13K14.90.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;

Bacteria; Magnoliophyta; eudicotyledons; core eudicots; rosids;
euroids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083488; PubMed=10617198;
RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
RA Harris B., Ansoorge W., Brandt P., Grivell L.A., Rieger M., Mueller M.,
RA Weichselgartner M., de Simone V., Obermaier B., Mache R., Schmidtheini T.,
RA Kreis M., Delseny M., Puigdomenech P., Watson M., Schimidtheini T.,
RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P.,
RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,
RA Braeken M., Welljens I., Voet M., Bastiaens I., Aert R., Braum E.,
RA Weitzenecker T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
RA Holzner E., Brandt A., Peters S., van Staveren M., Dirkse W.,
RA Molijman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
RA Berneiser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
RA De Keyser A., Buyshaert C., Gielen J., Villarroel R., De Clercq R.,
RA Van Montagu M., Rogers J., Cronin A., Quail M.A., Bray-Allen S.,
RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McIay K., Mayes R.,
RA Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
RA Borkova D., Bloecker H., Scharfe M., Grimm M., Loehnert T.-H.,
RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fartmann B., Grandrath K., Dauner D., Herzl A.,
RA Neumann S., Argirou A., Vitale D., Liguori R., Faravanti E.,
RA Massenon O., Quigley F., Clabaud G., Muendlein A., Felber R.,
RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA Chefdr F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA Gibbons T., Weber N., Vandenbol M., Barques M., Terol J., Torres A.,
RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielke C.,
RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.W.,
RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshua C.,
RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA Ma P., Zhong J., Preston R., Vil D., Shekhar M., Matero A., Shah R.,
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Tili S.,
RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
RA Chen E., Marra M.A., Martienssen R., McCombie W.R.;
RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
RT thaliana.";
RL Nature 402:769-777(1999).
RP SEQUENCE FROM N.A.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.A.;
RT "Full-length cDNA from Arabidopsis thaliana.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: 3-hydroxy-2-methylpropanoate + NAD(+) = 2-
CC methyl-3-oxopropanoate + NADH.
CC -!- SUBCELLULAR LOCATION: Mitochondrial (By similarity).
CC -!- SIMILARITY: Belongs to the 3-hydroxyisobutyrate dehydrogenase
CC family.
CC -!- CAUTION: Ref.1 sequences differ from that shown due to erroneous
CC gene model prediction.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

or send an email to license@isb-sib.ch).

CC EMBL; AL080282; CAB45888.1; ALT SEQ.
CC EMBL; ALI61554; CAB79093.1; ALT_SEQ.
DR EMBL; AY086845; AAM63893.1; -
DR InterPro: IPR002204; 3hydroxisobut dh.
DR PROSITE; PS00895; 3_HYDROXYISOBUT_DH; 1.
KW Hypothetical protein; Oxidoreductase; NAD; Mitochondrion;
KW Transit peptide.

KW TRANSIT 1 34 MITOCHONDRION (BY SIMILARITY).
FT CHAIN 35 347 PROBABLE 3-HYDROXYISOBUTYRATE
FT FT DEHYDROGENASE.
FT NP_BIND 38 66 NAD (ADP PART) (POTENTIAL).
FT ACT_SITE 219 219 BY SIMILARITY.
FT CONFLICT 21 21 S -> F (IN REF. 2).
SQ SEQUENCE 347 AA; 37364 MW; 6ED2E87CC3DE191C CRC64;

Query Match 37.5%; Score 75; DB 1; Length 347;
Best Local Similarity 39.5%; Pred.No. 0.016;
Matches 15; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

QY 1 GFGLGMLMGSGIVSNLLKMGHTVTWNRTAEKCDLFQI 38
||:||||| : ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 41 GFGLGWMGFVRVNNLRAGYKVTVHDIRVMKMFE 78

RESULT 24
6PGD STAEP STANDARD; PRT; 468 AA.
AC O8CP47;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE 6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44).
GN GND OR SEL192.
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12228;
RX PubMed=12950922;
RA Zhang Y.-Q., Ren S.-X., Li H.-L., Wang Y.-X., Fu G., Yang J.,
Qin Z.-Q., Miao Y.-G., Wang W.-Y., Chen R.-S., Shen Y., Chen Z.,
Guo Z.-H., Zhao G.-P., Qu D., Danchin A., Wen Y.-M.;
RT "Genome-based analysis of virulence genes in a non-biofilm-forming
Staphylococcus epidermidis strain (ATCC 12228)".
RM Mol. Microbiol. 49:1577-1593 (2003).
RL CC
CC -!- CATALYTIC ACTIVITY: 6-phospho-D-gluconate + NADP(+) = D-ribulose
5-phosphate + CO(2) + NADPH.
CC -!- PATHWAY: Hexose monophosphate shunt.
CC -!- SIMILARITY: Belongs to the 6-phosphogluconate dehydrogenase family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL; AE016748; AAC04791.1; -
DR InterPro: IPR008927; 6GDGH_C-like.
DR InterPro: IPR006183; 6PGD.
DR InterPro: IPR006184; 6PGdom_BS.
DR InterPro: IPR006114; 6PGD C.
DR InterPro: IPR006113; 6PGD decarbox.
DR InterPro: IPR006115; 6PGD_NAD.
DR Pfam; PF00393; 6PGD; 1.
DR Pfam; PF03446; NAD binding_2; 1.
DR PRINTS; PR00076; 6PGDHDRGNASE.
DR TIGRFAMS; TIGR00873; gnd; 1.
DR PROSITE; PS00461; 6PGD; 1.
KW Oxidoreductase; Pentose shunt; NADP.
SQ SEQUENCE 485 AA; 53102 MW; F7E6DDFF3D21EFB2 CRC64;

Query Match 37.5%; Score 75; DB 1; Length 485;
Best Local Similarity 37.8%; Pred.No. 0.023;
Matches 14; Conservative 7; Mismatches 16; Indels 0; Gaps 0;

QY 1 GFGLGMLMGSGIVSNLLKMGHTVTWNRTAEKCDLFQI 37
||:||||| : ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 9 GLIGLAVMGQNLIINMDHGFWVCAYNNTTSKVDLFL 45

or send an email to license@isb-sib.ch).

CC EMBL; AL080282; CAB45888.1; ALT SEQ.
CC EMBL; ALI61554; CAB79093.1; ALT_SEQ.
DR EMBL; AY086845; AAM63893.1; -
DR InterPro: IPR002204; 3hydroxisobut dh.
DR PROSITE; PS00895; 3_HYDROXYISOBUT_DH; 1.
KW Hypothetical protein; Oxidoreductase; NAD; Mitochondrion;
KW Transit peptide.

KW TRANSIT 1 34 MITOCHONDRION (BY SIMILARITY).
FT CHAIN 35 347 PROBABLE 3-HYDROXYISOBUTYRATE
FT FT DEHYDROGENASE.
FT NP_BIND 38 66 NAD (ADP PART) (POTENTIAL).
FT ACT_SITE 219 219 BY SIMILARITY.
FT CONFLICT 21 21 S -> F (IN REF. 2).
SQ SEQUENCE 347 AA; 37364 MW; 6ED2E87CC3DE191C CRC64;

Query Match 37.5%; Score 75; DB 1; Length 347;
Best Local Similarity 39.5%; Pred.No. 0.016;
Matches 15; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

QY 1 GFGLGMLMGSGIVSNLLKMGHTVTWNRTAEKCDLFQI 38
||:||||| : ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 41 GFGLGWMGFVRVNNLRAGYKVTVHDIRVMKMFE 78

RESULT 24
6PGD STAEP STANDARD; PRT; 468 AA.
AC O8CP47;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE 6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44).
GN GND OR SEL192.
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12228;
RX PubMed=12950922;
RA Zhang Y.-Q., Ren S.-X., Li H.-L., Wang Y.-X., Fu G., Yang J.,
Qin Z.-Q., Miao Y.-G., Wang W.-Y., Chen R.-S., Shen Y., Chen Z.,
Guo Z.-H., Zhao G.-P., Qu D., Danchin A., Wen Y.-M.;
RT "Genome-based analysis of virulence genes in a non-biofilm-forming
Staphylococcus epidermidis strain (ATCC 12228)".
RM Mol. Microbiol. 49:1577-1593 (2003).
RL CC
CC -!- CATALYTIC ACTIVITY: 6-phospho-D-gluconate + NADP(+) = D-ribulose
5-phosphate + CO(2) + NADPH.
CC -!- PATHWAY: Hexose monophosphate shunt.
CC -!- SIMILARITY: Belongs to the 6-phosphogluconate dehydrogenase family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL; AE016748; AAC04791.1; -
DR InterPro: IPR008927; 6GDGH_C-like.
DR InterPro: IPR006183; 6PGD.
DR InterPro: IPR006184; 6PGdom_BS.
DR InterPro: IPR006114; 6PGD C.
DR InterPro: IPR006113; 6PGD decarbox.
DR InterPro: IPR006115; 6PGD_NAD.
DR Pfam; PF00393; 6PGD; 1.
DR Pfam; PF03446; NAD binding_2; 1.
DR PRINTS; PR00076; 6PGDHDRGNASE.
DR TIGRFAMS; TIGR00873; gnd; 1.
KW Oxidoreductase; Pentose shunt; NADP.
SQ SEQUENCE 485 AA; 53102 MW; F7E6DDFF3D21EFB2 CRC64;

Query Match 37.5%; Score 75; DB 1; Length 485;
Best Local Similarity 37.8%; Pred.No. 0.023;
Matches 14; Conservative 7; Mismatches 16; Indels 0; Gaps 0;

QY 1 GFGLGMLMGSGIVSNLLKMGHTVTWNRTAEKCDLFQI 37
||:||||| : ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 9 GLIGLAVMGQNLIINMDHGFWVCAYNNTTSKVDLFL 45


```
RESULT 26
6PG9_ECOLI STANDARD; PRT; 468 AA.
ID 6PG9_ECOLI
AC P37754;
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44).
GN GND.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O9:K30:H12 / E69;
RX MEDLINE=94252978; PubMed=7515042;
RA Jayaratne P., Bronner D., MacLachlan R.P., Dodgson C., Kido N.,
RA Witfield C.;
RT Cloning and analysis of duplicated rfbM and rfbK genes involved in
RT the formation of GDP-mannose in Escherichia coli O9:K30 and
RT participation of rfb genes in the synthesis of the group I K30
RT capsular polysaccharide."
RL J. Bacteriol. 176:3126-3139(1994).
CC -1- CATALYTIC ACTIVITY: 6-phospho-D-gluconate + NADP(+) = D-ribulose
CC 5-phosphate + CO(2) + NADPH.
CC -1- PATHWAY: Hexose monophosphate shunt.
CC -1- SIMILARITY: Belongs to the 6-phosphogluconate dehydrogenase
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; L27646; AAA21136.1; -.
CC PIR; I41250; I41250.
CC DR HSP; P00349; 2PGD.
CC DR InterPro; IPR008927; 6GDGH_C_like.
CC DR InterPro; IPR006183; 6PGD.
CC DR InterPro; IPR006114; 6PGD_C.
CC DR InterPro; IPR006114; 6PGD_C.
CC DR InterPro; IPR006113; 6PGD decarbox.
CC DR InterPro; IPR006115; 6PGD_NAD.
CC DR Pfam; PF00393; 6PGD; 1.
CC DR PRINTS; PR00076; 6PGDHRGNASE.
CC DR TIGRfams; TIGR00873; gnd; 1.
CC DR PROSITE; PS00461; 6PGD; 1.
CC DR PROSITE; PS00461; 6PGD; 1.
CC KW Gluconate utilization; Oxidoreductase; Pentose shunt; NADP.
CC Gluconate utilization; Oxidoreductase; Pentose shunt; NADP.
CC CONFLICT 316 316 G -> E (IN REF. 2).
CC CONFLICT 421 421 V -> F (IN REF. 2).
CC SEQUENCE 468 AA; 51625 MW; C13D94CFD78BFF3A CRC64;
SQ
Query Match 37.08; Score 74; DB 1; Length 468;
Best Local Similarity 38.24; Pred. No. 0.03;
Matches 13; Conservative 10; Mismatches 11; Indels 0; Gaps 0;
QY 1 GFLGLGLMGSGIVSNLLKMGHTVTVMNRKAEKCD 34
| : : : : : : : : : : : : : : : : : :
Db 7 GVVGMAVGNRLALNIESRGYTVSVNRSREKTE 40

RESULT 27
6PGD_KLEPN STANDARD; PRT; 468 AA.
ID 6PGD_KLEPN
AC P41576; O48461;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44).
GN GND.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O9:K30:H12 / E69;
RX MEDLINE=94252978; PubMed=7515042;
RA Jayaratne P., Bronner D., MacLachlan R.P., Dodgson C., Kido N.,
RA Witfield C.;
RT Cloning and analysis of duplicated rfbM and rfbK genes involved in
RT the formation of GDP-mannose in Escherichia coli O9:K30 and
RT participation of rfb genes in the synthesis of the group I K30
RT capsular polysaccharide."
RL J. Bacteriol. 176:3126-3139(1994).
CC -1- CATALYTIC ACTIVITY: 6-phospho-D-gluconate + NADP(+) = D-ribulose
CC 5-phosphate + CO(2) + NADPH.
CC -1- PATHWAY: Hexose monophosphate shunt.
CC -1- SIMILARITY: Belongs to the 6-phosphogluconate dehydrogenase
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; L27646; AAA21136.1; -.
CC PIR; I41250; I41250.
CC DR HSP; P00349; 2PGD.
CC DR InterPro; IPR008927; 6GDGH_C_like.
CC DR InterPro; IPR006183; 6PGD.
CC DR InterPro; IPR006114; 6PGD_C.
CC DR InterPro; IPR006114; 6PGD_C.
CC DR InterPro; IPR006113; 6PGD decarbox.
CC DR InterPro; IPR006115; 6PGD_NAD.
CC DR Pfam; PF00393; 6PGD; 1.
CC DR PRINTS; PR00076; 6PGDHRGNASE.
CC DR TIGRfams; TIGR00873; gnd; 1.
CC DR PROSITE; PS00461; 6PGD; 1.
CC DR PROSITE; PS00461; 6PGD; 1.
CC KW Gluconate utilization; Oxidoreductase; Pentose shunt; NADP.
CC Gluconate utilization; Oxidoreductase; Pentose shunt; NADP.
CC CONFLICT 316 316 G -> E (IN REF. 2).
CC CONFLICT 421 421 V -> F (IN REF. 2).
CC SEQUENCE 468 AA; 51625 MW; C13D94CFD78BFF3A CRC64;
SQ
Query Match 37.08; Score 74; DB 1; Length 468;
Best Local Similarity 38.24; Pred. No. 0.03;
Matches 13; Conservative 10; Mismatches 11; Indels 0; Gaps 0;
QY 1 GFLGLGLMGSGIVSNLLKMGHTVTVMNRKAEKCD 34
| : : : : : : : : : : : : : : : : : :
Db 7 GVVGMAVGNRLALNIESRGYTVSVNRSREKTE 40

RESULT 28
6PGD_SALTY STANDARD; PRT; 468 AA.
ID 6PGD_SALTY
AC P14062;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DE 6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44).
GN GND.
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Klebsiella.
OX NCBI_TaxID=573;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Chedd;
RX MEDLINE=95204345; PubMed=7896702;
RA Arakawa Y., Wacharotayankun R., Nagatsuka T., Ito H., Kato N.,
RA Ohta M.;
RT "Genomic organization of the Klebsiella pneumoniae cps region
RT responsible for serotype K2 capsular polysaccharide synthesis in the
RT virulent strain Chedd."
RL J. Bacteriol. 177:1788-1796(1995).
RN [2]
RP SEQUENCE OF 12-456 FROM N.A.
RC STRAIN=CW 7380;
RX MEDLINE=95024018; PubMed=7937867;
RA Nelson K., Selander R.K.;
RT Intergenic transfer and recombination of the 6-phosphogluconate
RT dehydrogenase gene (gnd) in enteric bacteria."
RL Proc. Natl. Acad. Sci. U.S.A. 91:10227-10231(1994).
CC -1- CATALYTIC ACTIVITY: 6-phospho-D-gluconate + NADP(+) = D-ribulose
CC 5-phosphate + CO(2) + NADPH.
CC -1- PATHWAY: Hexose monophosphate shunt.
CC -1- SIMILARITY: Belongs to the 6-phosphogluconate dehydrogenase
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; D21242; BAA04786.1; -.
CC PIR; D56146; D56146.
CC DR HSP; P00349; 2PGD.
CC DR InterPro; IPR008927; 6GDGH_C_like.
CC DR InterPro; IPR006183; 6PGD.
CC DR InterPro; IPR006114; 6PGD_C.
CC DR InterPro; IPR006114; 6PGD_C.
CC DR InterPro; IPR006113; 6PGD decarbox.
CC DR InterPro; IPR006115; 6PGD_NAD.
CC DR Pfam; PF00393; 6PGD; 1.
CC DR PRINTS; PR00076; 6PGDHRGNASE.
CC DR TIGRfams; TIGR00873; gnd; 1.
CC DR PROSITE; PS00461; 6PGD; 1.
CC DR PROSITE; PS00461; 6PGD; 1.
CC KW Gluconate utilization; Oxidoreductase; Pentose shunt; NADP.
CC Gluconate utilization; Oxidoreductase; Pentose shunt; NADP.
CC CONFLICT 316 316 G -> E (IN REF. 2).
CC CONFLICT 421 421 V -> F (IN REF. 2).
CC SEQUENCE 468 AA; 51328 MW; FF1EB5E765FDC90 CRC64;
SQ
Query Match 37.08; Score 74; DB 1; Length 468;
Best Local Similarity 38.24; Pred. No. 0.03;
Matches 13; Conservative 10; Mismatches 11; Indels 0; Gaps 0;
QY 1 GFLGLGLMGSGIVSNLLKMGHTVTVMNRKAEKCD 34
| : : : : : : : : : : : : : : : : : :
Db 7 GVVGMAVGNRLALNIESRGYTVSVNRSREKTE 40

RESULT 29
6PGD_KLEPN STANDARD; PRT; 468 AA.
ID 6PGD_KLEPN
AC P41576; O48461;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44).
GN GND.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O9:K30:H12 / E69;
RX MEDLINE=94252978; PubMed=7515042;
RA Jayaratne P., Bronner D., MacLachlan R.P., Dodgson C., Kido N.,
RA Witfield C.;
RT Cloning and analysis of duplicated rfbM and rfbK genes involved in
RT the formation of GDP-mannose in Escherichia coli O9:K30 and
RT participation of rfb genes in the synthesis of the group I K30
RT capsular polysaccharide."
RL J. Bacteriol. 176:3126-3139(1994).
CC -1- CATALYTIC ACTIVITY: 6-phospho-D-gluconate + NADP(+) = D-ribulose
CC 5-phosphate + CO(2) + NADPH.
CC -1- PATHWAY: Hexose monophosphate shunt.
CC -1- SIMILARITY: Belongs to the 6-phosphogluconate dehydrogenase
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; L27646; AAA21136.1; -.
CC PIR; I41250; I41250.
CC DR HSP; P00349; 2PGD.
CC DR InterPro; IPR008927; 6GDGH_C_like.
CC DR InterPro; IPR006183; 6PGD.
CC DR InterPro; IPR006114; 6PGD_C.
CC DR InterPro; IPR006114; 6PGD_C.
CC DR InterPro; IPR006113; 6PGD decarbox.
CC DR InterPro; IPR006115; 6PGD_NAD.
CC DR Pfam; PF00393; 6PGD; 1.
CC DR PRINTS; PR00076; 6PGDHRGNASE.
CC DR TIGRfams; TIGR00873; gnd; 1.
CC DR PROSITE; PS00461; 6PGD; 1.
CC DR PROSITE; PS00461; 6PGD; 1.
CC KW Gluconate utilization; Oxidoreductase; Pentose shunt; NADP.
CC Gluconate utilization; Oxidoreductase; Pentose shunt; NADP.
CC CONFLICT 316 316 G -> E (IN REF. 2).
CC CONFLICT 421 421 V -> F (IN REF. 2).
CC SEQUENCE 468 AA; 51625 MW; C13D94CFD78BFF3A CRC64;
SQ
Query Match 37.08; Score 74; DB 1; Length 468;
Best Local Similarity 38.24; Pred. No. 0.03;
Matches 13; Conservative 10; Mismatches 11; Indels 0; Gaps 0;
QY 1 GFLGLGLMGSGIVSNLLKMGHTVTVMNRKAEKCD 34
| : : : : : : : : : : : : : : : : : :
Db 7 GVVGMAVGNRLALNIESRGYTVSVNRSREKTE 40
```


28-FEB-2003 (Rel. 41, Last annotation update)
6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44).
GND OR STM2081.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RX MEDLINE=89364685; PubMed=2671649;
RA Reeves P., Stevenson G.;
RT "Cloning and nucleotide sequence of the Salmonella typhimurium LT2
gnd gene and its homology with the corresponding sequence of
Escherichia coli K12.";
RL Mol. Gen. Genet. 217:182-184 (1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RX Dykhuizen D.E., Green L.;
RA Submitted (XXX-1991) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanders K.E., Spieth J., Clifton S.W., Latreille P.,
Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
Waterston R., Wilson R.K.;
RA "Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2.";
RL Nature 413:852-856 (2001).
RN [4]
RP SEQUENCE OF 1-57 FROM N.A.
RC STRAIN=LT2;
RX MEDLINE=91260454; PubMed=1710759;
RA Jiang X.M., Neal B., Santiago F., Lee S.J., Romana L.K., Reeves P.R.;
RT "Structure and sequence of the rfb (O antigen) gene cluster of
Salmonella serovar typhimurium (strain LT2).";
RL Mol. Microbiol. 5:695-713 (1991).
CC -1- CATALYTIC ACTIVITY: 6-phospho-D-gluconate + NADP(+) = D-ribulose
5-phosphate + CO(2) + NADPH.
CC -1- PATHWAY: Hexose monophosphate shunt.
CC -1- SIMILARITY: Belongs to the 6-phosphogluconate dehydrogenase
family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X15651; CAA33677.1; -.
CC EMBL; MG4332; AAA27137.1; -.
CC EMBL; AE008792; AAL20985.1; -.
CC EMBL; X56793; CAA40131.1; -.
CC F1R; S04397; S04397.
CC HSSP; P00349; 2PGD.
CC StyGene; SG10146; gnd.
CC InterPro; IPR008927; 6GDGH_C_like.
CC InterPro; IPR006183; 6PGD.
CC InterPro; IPR006114; 6PGD_C.
CC InterPro; IPR006113; 6PGD Decarbox.
CC InterPro; IPR006115; 6PGD NAD.
CC InterPro; IPR006184; 6PGDom_BS.
CC Pfam; PF00393; 6PGD; 1.
CC PRINTS; PR00076; NAD_binding_2; 1.
CC TIGRfam; TIGR00873; gnd; 1.
CC PROSITE; PS00461; 6PGD; 1.
CC

KW Gluconate utilization; Oxidoreductase; Pentose shunt; NADP;
KW Complete proteome.
SQ SEQUENCE 468 AA; 51395 MW; D8EB53A2DAADB7 CRC64;
Query Match 37.0%; Score 74; DB 1; Length 468;
Best Local Similarity 38.2%; Pred. No. 0.03;
Matches 13; Conservative 10; Mismatches 11; Indels 0; Gaps 0;
QY 1 GFLGLMGSGIVSNLLKMGHTVTVMNRTEAKCD 34
Db 7 GVGMAYMGRNALNIESRGYTVSVFNRSEKTE 40
RESULT 29
6PGD_CERCA STANDARD; PRT; 481 AA.
ID -6PGD_CERCA
AC P41570;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44).
GN PGD.
OS Ceratitidis capitata (Mediterranean fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Tephritidae; Tephritidae; Ceratitidis.
OX NCBI_TaxID=7213;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94093871; PubMed=8269100;
RA Scott M.J., Kriticou D., Robinson A.S.;
RT "Isolation of cDNAs encoding 6-phosphogluconate dehydrogenase and
glucose-6-phosphate dehydrogenase from the Mediterranean fruit fly
Ceratitidis capitata: correlating genetic and physical maps of
chromosome 5.";
RL Insect Mol. Biol. 1:213-222 (1993).
CC -1- CATALYTIC ACTIVITY: 6-phospho-D-gluconate + NADP(+) = D-ribulose
5-phosphate + CO(2) + NADPH.
CC -1- PATHWAY: Hexose monophosphate shunt.
CC -1- SIMILARITY: Belongs to the 6-phosphogluconate dehydrogenase
family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; S67873; AAB29396.1; -.
CC HSSP; P00349; 2PGD.
CC InterPro; IPR008927; 6GDGH_C_like.
CC InterPro; IPR006183; 6PGD.
CC InterPro; IPR006114; 6PGD_C.
CC InterPro; IPR006113; 6PGD Decarbox.
CC InterPro; IPR006115; 6PGD NAD.
CC InterPro; IPR006184; 6PGDom_BS.
CC Pfam; PF00393; 6PGD; 1.
CC Pfam; PF03446; NAD_binding_2; 1.
CC PRINTS; PR00076; 6PGDHRGNASE.
CC TIGRfam; TIGR00873; gnd; 1.
CC PROSITE; PS00461; 6PGD; 1.
KW Gluconate utilization; Oxidoreductase; Pentose shunt; NADP.
SQ SEQUENCE 481 AA; 52963 MW; F0ABB506AD1B86D0 CRC64;
Query Match 37.0%; Score 74; DB 1; Length 481;
Best Local Similarity 38.9%; Pred. No. 0.03;
Matches 14; Conservative 8; Mismatches 14; Indels 0; Gaps 0;
QY 3 LGGLMGSGIVSNLLKMGHTVTVMNRTEAKCDLFIQ 38
Db 10 IGLAVMGQNLVLNMNDKGFVVCAYNFTVEKVNQFLK 45

RESULT 30

6PGD-SCHPO STANDARD; PRT; 492 AA.

ID 6PGD-SCHPO STANDARD; PRT; 492 AA.

AC P78812; Q9UQW5;

DT 15-JUL-1998 (Rel. 36, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE 6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44).

GN SPBC660.16.

OS Schizosaccharomyces pombe (Fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomycetes.

OX NCBI_TaxID=4896;

[1]

RP SEQUENCE FROM N.A.

RP STRAIN=PR745;

RX MEDLINE=98162722; PubMed=9501991;

RY Yoshioka S., Kato K., Nakai K., Okayama H., Nojima H.;

RT "Identification of open reading frames in Schizosaccharomyces pombe

RT CDNAS.";

RL DNA Res. 4:363-369(1997).

[2]

RP SEQUENCE FROM N.A.

RP STRAIN=972;

RX MEDLINE=21848401; PubMed=11859360;

RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,

RA Scours J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,

RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,

RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,

RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,

RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,

RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,

RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeil C.,

RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,

RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,

RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,

RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,

RA Woodward J., Volckreels G., Aert R., Robben J., Grymonprez B.,

RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,

RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,

RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,

RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,

RA Goffeau A., Cadieu E., Drenth S., Gloux S., Lelaure V., Mottier S.,

RA Galibert F., Ades S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,

RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,

RA Daga R.R., Cruzado I., Jimenez J., Sanchez M., del Rey F., Benito J.,

RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,

RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,

RA Shpakowski G.V., Ussery D., Barrall B.G., Nurse P.;

RT "The genome sequence of Schizosaccharomyces pombe.";

RL Nature 415:871-880(2002).

CC -!- CATALYTIC ACTIVITY: 6-phospho-D-gluconate + NADP(+) = D-ribulose

CC 5-phosphate + CO(2) + NADPH.

CC -!- PATHWAY: Hexose monophosphate shunt.

CC -!- SIMILARITY: Belongs to the 6-phosphogluconate dehydrogenase

CC family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

CC or send an email to license@sb-sib.ch).

DR EMBL; D89161; BAAL3823.1; ALT_INIT.

DR EMBL; AL034563; CAA22536.1; -.

DR FIR; T40628; T40628.

DR HSSP; P00349; 2PGD.

DR GeneDB SPombe; SPBC660.16; -.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2004, 07:30:52 ; Search time 39 Seconds
(without alignments)
307.428 Million cell updates/sec

Title: US-10-067-482-2_COPY_271_308

Perfect score: 200

Sequence: 1 GFLGLMGSGIVSNLLMGMHTVTVWARTAEKCDLFIQ 38

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

- SPTREMBL 25:*
- 1: sp_archea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phase:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*
 - 15: sp_rvirus:*
 - 16: sp_bacteriap:*
 - 17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	200	100.0	523	4	Q9BT11
2	166	83.0	546	11	Q922P9
3	166	83.0	546	11	Q9CYQ1
4	166	83.0	547	4	Q9EXK2
5	126	63.0	602	5	Q8T079
6	125	62.5	248	10	Q84A74
7	125	62.5	289	10	Q94B07
8	125	62.5	289	10	Q9LSV0
9	123	61.5	343	10	Q8RWF1
10	123	61.5	387	10	Q9LNP0
11	121	60.5	343	10	Q8LQJ7
12	118	59.0	293	10	Q94VC8
13	109	54.5	292	16	Q97DX8
14	108	54.0	289	17	Q97XZ7
15	105	52.5	155	16	Q92N92
16	101	50.5	300	17	Q8TT25
					Q9bt11 homo sapien
					Q922p9 mus musculus
					Q9cyq1 mus musculus
					Q9bxx2 homo sapien
					Q8t079 drosophila
					Q94a74 arabidopsis
					Q94b07 arabidopsis
					Q9lsv0 arabidopsis
					Q8rwl1 arabidopsis
					Q9lnp0 arabidopsis
					Q8lqj7 oryza sativ
					Q94vc8 oryza sativ
					Q97dx8 clostridium
					Q97xz7 sulfolobus
					Q92n92 rhizobium m
					Q8tt25 methanosarc

17	100	50.0	286	16	Q88UE9
18	99	49.5	292	16	Q819E3
19	98	49.0	291	16	Q911R8
20	96	48.0	295	16	Q88NR2
21	96	48.0	313	16	Q89HA0
22	96	48.0	317	16	Q81S35
23	96	48.0	475	16	Q88GH7
24	93	46.5	296	16	Q82MM0
25	93	46.5	290	16	Q8FT11
26	92	46.0	43	2	Q05188
27	92	46.0	301	16	Q7WFP7
28	92	46.0	302	16	Q7W489
29	92	46.0	334	10	Q9S2E1
30	91	45.5	292	16	Q8PQK1
31	91	45.5	493	5	Q8TA03
32	90	45.0	293	16	Q930R8
33	90	45.0	298	16	Q7WJP8
34	90	45.0	298	16	Q7W4J7
35	89	44.5	291	16	Q8EDH8
36	89	44.5	296	16	Q830A7
37	89	44.5	469	16	Q88RQ4
38	88	44.0	285	16	Q7U3K2
39	88	44.0	291	16	Q8PDM7
40	88	44.0	293	16	Q89M84
41	88	44.0	296	16	Q9RD04
42	88	44.0	464	16	Q51509
43	88	44.0	468	16	Q99TY2
44	88	44.0	469	16	Q9WYR9
45	87	43.5	291	2	Q9RBX6
46	87	43.5	345	5	Q86R98
47	87	43.5	468	16	Q931R3
48	86.5	43.2	290	16	Q89NU4
49	86	43.0	286	16	Q92D17
50	86	43.0	298	16	Q8ESP3
51	86	43.0	329	16	Q8D6N9
52	86	43.0	472	16	Q92BX9
53	86	43.0	472	16	Q8Y7B0
54	86	43.0	491	3	Q8NKA5
55	86	43.0	482	16	Q8NQ12
56	85	42.5	481	16	Q823H8
57	84.5	42.2	296	16	Q8ZLV8
58	84.5	42.2	299	16	Q8XAE4
59	84.5	42.2	299	16	Q83Q26
60	84	42.0	288	16	Q66454
61	84	42.0	295	16	Q92TT0
62	84	42.0	299	10	Q9C991
63	83	41.5	54	16	Q89BE7
64	83	41.5	311	16	Q9L2D7
65	83	41.5	467	16	Q8RD52
66	83	41.5	491	3	Q874Q3
67	82.5	41.2	291	16	Q81GY1
68	82.5	41.2	291	16	Q8G0Y2
69	82.5	41.2	294	16	Q823K1
70	82	41.0	293	16	Q9CDK4
71	82	41.0	297	10	Q9C990
72	82	41.0	300	16	Q8UBW3
73	82	41.0	305	16	Q82Q44
74	82	41.0	318	10	Q949M8
75	81	40.5	286	16	Q8Y8A5
76	81	40.5	458	5	Q8SRX1
77	81	40.5	476	16	Q8UF73
78	81	40.5	492	3	Q875M5
79	81	40.5	493	3	Q876C3
80	80	40.0	261	3	Q876H8
81	80	40.0	289	17	Q97ZE5
82	80	40.0	298	16	Q8X8D4
83	80	40.0	298	16	Q8Z2T6
84	80	40.0	316	10	Q8LQ70
85	80	40.0	474	16	Q7VRX5
86	80	40.0	476	16	Q8YLM3
87	80	40.0	479	5	Q9NGR0
88	80	40.0	479	5	Q9NGR1
89	80	40.0	479	5	Q9SUN7

Q88ue9 lactobacill
Q819e3 bacillus ce
Q911r8 pseudomonas
Q88nr2 pseudomonas
Q89ha0 bradyrhizob
Q81s35 bacillus an
Q88gh7 rhizobium l
Q82mw0 streptomyce
Q8ft11 corynebacte
Q05188 bacillus su
Q7wfp7 bordetella
Q7w489 bordetella
Q9s2e1 arabidopsis
Q8pzk1 xanthomonas
Q8ta03 dictyosteli
Q930r8 rhizobium m
Q7wjp8 bordetella
Q7w4j7 bordetella
Q8edh8 shewanella
Q830a7 enterococcu
Q8r8q4 thermoaer
Q7u3k2 synchococc
Q8pdm7 xanthomonas
Q89m84 bradyrhizob
Q8rd04 streptomyce
Q51509 borrelia bu
Q99ty2 staphylococ
Q9wy9 thermotoga
Q8rbx6 pseudomonas
Q86r98 drosophila
Q931r3 staphylococ
Q89nu4 bradyrhizob
Q92d17 listeria in
Q8esp3 oceanobacil
Q8d6n9 vibrio vuln
Q92bx9 listeria in
Q8y7b0 listeria mo
Q8nka5 aspergillus
Q8nq12 corynebacte
Q823h8 chlamydophi
Q8zlv8 salmonella
Q8xae4 escherichia
Q83q26 shigella fl
Q66454 aquifex aeo
Q92tt0 rhizobium m
Q9c991 arabidopsis
Q89be7 bradyrhizob
Q9l2d7 streptomyce
Q8rd52 thermoaer
Q874q3 aspergillus
Q8ygy1 bruceella me
Q8g0y2 bruceella su
Q823k1 salmonella
Q9cdk4 lactococcus
Q9c990 arabidopsis
Q8ubw3 agrobacteri
Q82q44 streptomyce
Q949m8 arabidopsis
Q8y8a5 listeria mo
Q8srx1 encephalito
Q8uf73 agrobacteri
Q875m5 kluyveromyc
Q876c3 saccharomyc
Q876h8 saccharomyc
Q97ze5 sulfolobus
Q8x8d4 escherichia
Q8z2t6 salmonella
Q8lq70 oryza sativ
Q7vrx5 candidatus
Q8ylm3 anabaena sp
Q9ngr0 leishmania
Q9ngr1 leishmania
Q9sun7 leishmania

Q8VIV7 mycobacteri
P95165 mycobacteri
Q94KU2 spinacia ol
Q92PA2 rhizobium m
Q8FBG0 escherichia
Q8VAX7 brucella me
Q8FXG6 brucella su
Q92PE1 rhizobium m
Q9CKG2 pasteurella
Q7VMX4 haemophilus
Q876H9 saccharomyc

ALIGNMENTS

RESULT 1
Q9BT11 Q9BT11 PRELIMINARY; PRT; 523 AA.
AC Q9BT11; STRAUSBERG R.;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to RIKEN cDNA 3930401K13 gene (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003693; AA03693.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004616; F:phosphogluconate dehydrogenase (decarboxyla. . .; IEA.
DR GO; GO:0006098; P:pentose-phosphate shunt; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR006115; 6PGD_NAD.
DR InterPro; IPR000313; PWWP_domain.
DR Pfam; PF02178; AT hook; 1.
DR Pfam; PF03446; NAD_binding_2; 1.
DR SMART; SM00855; PWWP; 1.
DR PROSITE; PS50812; PWWP; 1.
FT NON_TER 1
SQ SEQUENCE 523 AA; 57215 MW; ED3E2853938A211F CRC64;

Query Match 100.0%; Score 200; DB 4; Length 523;
Best Local Similarity 100.0%; Pred. No. 1.6e-17;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GFLGLGLMGSGIVSNLLKMGHTVTWNRTAEKCDLFIQ 38
DB 241 GFLGLGLMGSGIVSNLLKMGHTVTWNRTAEKCDLFIQ 278
|||||

RESULT 2
Q922P9 Q922P9 PRELIMINARY; PRT; 546 AA.
AC Q922P9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE RIKEN cDNA 3930401K13 gene.
GN NPAC OR 3930401K13RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003693; AA03693.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004616; F:phosphogluconate dehydrogenase (decarboxyla. . .; IEA.
DR GO; GO:0006098; P:pentose-phosphate shunt; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR006115; 6PGD_NAD.
DR InterPro; IPR000313; PWWP_domain.
DR Pfam; PF02178; AT hook; 1.
DR Pfam; PF03446; NAD_binding_2; 1.
DR Pfam; PF00855; PWWP; 1.
DR SMART; SM00855; AT hook; 1.
DR PROSITE; PS50812; PWWP; 1.
FT NON_TER 1
SQ SEQUENCE 523 AA; 57215 MW; ED3E2853938A211F CRC64;

Query Match 100.0%; Score 200; DB 4; Length 523;
Best Local Similarity 100.0%; Pred. No. 1.6e-17;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GFLGLGLMGSGIVSNLLKMGHTVTWNRTAEKCDLFIQ 38
DB 241 GFLGLGLMGSGIVSNLLKMGHTVTWNRTAEKCDLFIQ 278
|||||

RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC006893; AA06893.1; -.
DR MGD; MGI:1921272; Npac.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004616; F:phosphogluconate dehydrogenase (decarboxyla. . .; IEA.
DR GO; GO:0006098; P:pentose-phosphate shunt; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR006115; 6PGD_NAD.
DR InterPro; IPR000313; PWWP_domain.
DR Pfam; PF02178; AT hook; 1.
DR Pfam; PF03446; NAD_binding_2; 1.
DR Pfam; PF00855; PWWP; 1.
DR SMART; SM00855; PWWP; 1.
DR PROSITE; PS50812; PWWP; 1.
SQ SEQUENCE 546 AA; 59715 MW; F5D2090DE1F64723 CRC64;

Query Match 83.0%; Score 166; DB 11; Length 546;
Best Local Similarity 100.0%; Pred. No. 4.3e-13;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GFLGLGLMGSGIVSNLLKMGHTVTWNRTAEK 32
DB 270 GFLGLGLMGSGIVSNLLKMGHTVTWNRTAEK 301
|||||

RESULT 3
Q9CYQ1 Q9CYQ1 PRELIMINARY; PRT; 546 AA.
AC Q9CYQ1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 3930401K13RIK protein.
GN NPAC OR 3930401K13RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yananaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Kuehl P., Stauber F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK014456; BAB29363.1; -.
DR MGD; MGI:1921272; Npac.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004616; F:phosphogluconate dehydrogenase (decarboxyla. . .; IEA.
DR GO; GO:0006098; P:pentose-phosphate shunt; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.

Query Match 100.0%; Score 166; DB 11; Length 546;
Best Local Similarity 100.0%; Pred. No. 4.3e-13;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GFLGLGLMGSGIVSNLLKMGHTVTWNRTAEK 32
DB 270 GFLGLGLMGSGIVSNLLKMGHTVTWNRTAEK 301
|||||

RESULT 3
Q9CYQ1 Q9CYQ1 PRELIMINARY; PRT; 546 AA.
AC Q9CYQ1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 3930401K13RIK protein.
GN NPAC OR 3930401K13RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yananaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Kuehl P., Stauber F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK014456; BAB29363.1; -.
DR MGD; MGI:1921272; Npac.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004616; F:phosphogluconate dehydrogenase (decarboxyla. . .; IEA.
DR GO; GO:0006098; P:pentose-phosphate shunt; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.

DR InterPro: IPR006115; 6PGD NAD.
 DR InterPro: IPR000637; AT hook.
 DR InterPro: IPR000313; PWWP_domain.
 DR Pfam: PF02178; AT_hook; 1.
 DR Pfam: PF03446; NAD_binding_2; 1.
 DR Pfam: PF00855; PWWP; 1.
 DR SMART: SM00384; AT_hook; 1.
 DR PROSITE: PS0812; PWWP; 1.
 SQ SEQUENCE 546 AA; 59744 MW; B3AC1562477ABCF CRC64;

Query Match 83.0%; Score 166; DB 11; Length 546;
 Best Local Similarity 100.0%; Pred. No. 4.3e-13;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GFLGLGLMGSGIVSNLLKMGHTVTVMNRTEAK 32
 DB 270 GFLGLGLMGSGIVSNLLKMGHTVTVMNRTEAK 301

RESULT 4
 Q9BXK2 PRELIMINARY; PRT; 547 AA.
 AC Q9BXK2; 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Cytokine-like nuclear factor n-pac.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA New L., Han J.;
 RT "A novel cytokine-like nuclear factor, n-pac";
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 RM EMBL; AF215966; AAK15524.1; -;
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0004616; F:phosphoglucanase dehydrogenase (decarboxyla. . .; IEA.
 DR GO; GO:0006098; P:ribose-phosphate shunt; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro: IPR006115; 6PGD NAD.
 DR InterPro: IPR000637; AT hook.
 DR InterPro: IPR000313; PWWP_domain.
 DR Pfam: PF02178; AT_hook; 1.
 DR Pfam: PF03446; NAD_binding_2; 1.
 DR Pfam: PF00855; PWWP; 1.
 DR SMART: SM00384; AT_hook; 1.
 DR PROSITE: PS0812; PWWP; 1.
 SQ SEQUENCE 547 AA; 59827 MW; C7D785CCBF83204A CRC64;

Query Match 83.0%; Score 166; DB 4; Length 547;
 Best Local Similarity 100.0%; Pred. No. 4.3e-13;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GFLGLGLMGSGIVSNLLKMGHTVTVMNRTEAK 32
 DB 271 GFLGLGLMGSGIVSNLLKMGHTVTVMNRTEAK 302

RESULT 5
 Q8T079 PRELIMINARY; PRT; 602 AA.
 AC Q8T079; Q9VU51;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE LD22344D (CG4747-PA).
 GN BEST:LD22483 OR BEST:LD29743 OR CG4747.
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
 RA Borkov D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Folsler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodager, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradscky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
 RA Tupy J.L., Bergman C.M., Berman B.P., Carlson J.W., Celniker S.E.,
 RA Clamp M.E., Drysdale R.A., Emmert D., Frise E., de Grey A.D.N.J.,
 RA Harris N.L., Kronmiller B., Marshall B., Millburn G.H., Richter J.,
 RA Russo S., Searle S.M.J., Smith E., Shu S., Smutniak F.,
 RA Whitfield E.J., Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J.,
 RA Lewis S.E.;
 RT "Annotation of Drosophila melanogaster genome";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA FlyBase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA FlyBase;
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 RM EMBL; AY069497; AAL39642.1; -;

EMBL; AE003627; AAF52846.3; --
 FlyBase; Fgn0043456; BEST:ID22483.
 GO; GO:0004616; F:phosphogluconate dehydrogenase (decarboxyla. . .; IEA.
 GO; GO:0006098; P:phosphogluconate dehydrogenase (decarboxyla. . .; IEA.
 InterPro; IPR006115; 6PGD NAD.
 InterPro; IPR000313; PWWP domain.
 Pfam; PF03446; NAD_binding_2; 1.
 Pfam; PF00855; PWWP; 1.
 SMART; SM00293; PWWP; 1.
 PROSITE; PS00812; PWWP; 1.
 SEQUENCE 602 AA; 65252 MW; E004EEC610C23625 CRC64;
 Query Match 63.0%; Score 126; DB 5; Length 602;
 Best Local Similarity 60.5%; Pred. No. 7.4e-08;
 Matches 23; Conservative 5; Mismatches 10; Indels 0; Gaps 0;
 QY 1 GFGLGLMGSGIVSNLLKMGHTVTWVNRTRAEKCDLFIQ 38
 DB 319 GFGLGLMGSGIVSNLLKMGHTVTWVNRTRAEKCDLFIQ 356
 RESULT 6
 Q94A74 PRELIMINARY; PRT; 248 AA.
 ID Q94A74
 AC Q94A74
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE AT325530/MWL2.15.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Kim C.J., Chen H., Cheuk R., Koesema E., Meyers M.C., Banh J.,
 RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
 RA Tshida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
 RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 RT "Arabidopsis cDNA clones.";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AV049298; AAK83640.1; --
 DR GO; GO:0008442; F:3-hydroxyisobutyrate dehydrogenase activity; IEA.
 DR GO; GO:0004616; F:phosphogluconate dehydrogenase (decarboxyla. . .; IEA.
 DR GO; GO:0006098; P:phosphogluconate dehydrogenase (decarboxyla. . .; IEA.
 DR GO; GO:0006573; P:valine metabolism; IEA.
 DR InterPro; IPR002204; 3hydroxisobut_dh.
 DR InterPro; IPR006183; 6PGD.
 DR Pfam; PF03446; NAD_binding_2; 1.
 DR PRINTS; PR00076; 6PGDHDRGNASE.
 DR PROSITE; PS00895; 3_HYDROXYISOBUT_DH; 1.
 DR SEQUENCE 248 AA; 26336 MW; C96912B3E2A7B677 CRC64;
 Query Match 62.5%; Score 125; DB 10; Length 248;
 Best Local Similarity 60.5%; Pred. No. 3.9e-08;
 Matches 23; Conservative 4; Mismatches 11; Indels 0; Gaps 0;
 QY 1 GFGLGLMGSGIVSNLLKMGHTVTWVNRTRAEKCDLFIQ 38
 DB 4 GFGLGLMGSGIVSNLLKMGHTVTWVNRTRAEKCDLFIQ 41
 RESULT 7
 Q94B07 PRELIMINARY; PRT; 289 AA.
 ID Q94B07
 AC Q94B07
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE AT325530/MWL2.15.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Kim C.J., Chen H., Cheuk R., Koesema E., Meyers M.C., Banh J.,
 RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
 RA Tshida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
 RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 RT "Arabidopsis cDNA clones.";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AV049298; AAK83640.1; --
 DR GO; GO:0008442; F:3-hydroxyisobutyrate dehydrogenase activity; IEA.
 DR GO; GO:0004616; F:phosphogluconate dehydrogenase (decarboxyla. . .; IEA.
 DR GO; GO:0006098; P:phosphogluconate dehydrogenase (decarboxyla. . .; IEA.
 DR GO; GO:0006573; P:valine metabolism; IEA.
 DR InterPro; IPR002204; 3hydroxisobut_dh.
 DR InterPro; IPR006183; 6PGD.
 DR InterPro; IPR006115; 6PGD NAD.
 DR Pfam; PF03446; NAD_binding_2; 1.
 DR PRINTS; PR00076; 6PGDHDRGNASE.
 DR PROSITE; PS00895; 3_HYDROXYISOBUT_DH; 1.
 DR SEQUENCE 248 AA; 26336 MW; C96912B3E2A7B677 CRC64;
 Query Match 62.5%; Score 125; DB 10; Length 248;
 Best Local Similarity 60.5%; Pred. No. 3.9e-08;
 Matches 23; Conservative 4; Mismatches 11; Indels 0; Gaps 0;
 QY 1 GFGLGLMGSGIVSNLLKMGHTVTWVNRTRAEKCDLFIQ 38
 DB 4 GFGLGLMGSGIVSNLLKMGHTVTWVNRTRAEKCDLFIQ 41
 RESULT 7
 Q94B07 PRELIMINARY; PRT; 289 AA.
 ID Q94B07
 AC Q94B07
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE AT325530/MWL2.15.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Kim C.J., Chen H., Cheuk R., Koesema E., Meyers M.C., Banh J.,
 RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
 RA Tshida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
 RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 RT "Arabidopsis cDNA clones.";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AV049298; AAK83640.1; --
 DR GO; GO:0008442; F:3-hydroxyisobutyrate dehydrogenase activity; IEA.
 DR GO; GO:0004616; F:phosphogluconate dehydrogenase (decarboxyla. . .; IEA.
 DR GO; GO:0006098; P:phosphogluconate dehydrogenase (decarboxyla. . .; IEA.
 DR GO; GO:0006573; P:valine metabolism; IEA.
 DR InterPro; IPR002204; 3hydroxisobut_dh.
 DR InterPro; IPR006183; 6PGD.
 DR InterPro; IPR006115; 6PGD NAD.
 DR Pfam; PF03446; NAD_binding_2; 1.
 DR PRINTS; PR00076; 6PGDHDRGNASE.
 DR PROSITE; PS00895; 3_HYDROXYISOBUT_DH; 1.
 DR SEQUENCE 248 AA; 26336 MW; C96912B3E2A7B677 CRC64;
 Query Match 62.5%; Score 125; DB 10; Length 248;
 Best Local Similarity 60.5%; Pred. No. 3.9e-08;
 Matches 23; Conservative 4; Mismatches 11; Indels 0; Gaps 0;
 QY 1 GFGLGLMGSGIVSNLLKMGHTVTWVNRTRAEKCDLFIQ 38
 DB 4 GFGLGLMGSGIVSNLLKMGHTVTWVNRTRAEKCDLFIQ 41
 RESULT 7
 Q94B07 PRELIMINARY; PRT; 289 AA.
 ID Q94B07
 AC Q94B07
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE AT325530/MWL2.15.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Kim C.J., Chen H., Cheuk R., Koesema E., Meyers M.C., Banh J.,
 RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
 RA Tshida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
 RA Nguyen

```

DR InterPro; IPR006115; 6PGD NAD.
DR Pfam; PF03446; NAD binding 2; 1.
DR PRINTS; PR00076; 6PGDHDRGNASE.
DR PROSITE; PS00895; 3_HYDROXYISOBUT_DH; 1.
SQ SEQUENCE 289 AA; 30691 MW; 3567002A19706636 CRC64;

Query Match 62.5%; Score 125; DB 10; Length 289;
Best Local Similarity 60.5%; Pred. No. 4.5e-08;
Matches 23; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

Qy 1 GFLGLGMLGSGIVSNLLKMGHTVTVWNRATKCDLFIQ 38
Dd 4 GFLGLGMLGSGIVSNLLKMGHTVTVWNRATKCDLFIQ 41

RESULT 9
ID Q8RWFL PRELIMINARY; PRT; 343 AA.
AC Q8RWFL;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Atg17650).
GN Atg17650.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RA Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Haas B.J., Volfvsky N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RA "Full-length messenger RNA sequences greatly improve genome
RA annotation.";
RN [3]
RP SEQUENCE FROM N.A.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RA Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Kim C.J., Chen H., Cheuk R., Shinn P., Bowser L., Carninci P.,
RA Dale J.M., Hayashizaki Y., Heuan V.W., Ishida J., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Southwick A., Toriumi M., Wong C.; Wu H.C., Yamada K., Yu G.,
RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RA "Arabidopsis ORF clones.";
RA Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
RN [5]
RP EMBL; AY093135; AAM13134.1; -.
DR EMBL; AY085690; AAM62909.1; -.
DR EMBL; BT008734; AAP42747.1; -.
DR GO; GO:0004616; F:phosphogluconate dehydrogenase (decarboxyla. . .; IEA.
DR GO; GO:0006098; P:pentose-phosphate shunt; IEA.
DR InterPro; IPR006115; 6PGD NAD.
DR Pfam; PF03446; NAD binding_2; 1.
DR Hypothetical protein.
SQ SEQUENCE 343 AA; 36287 MW; D012C16284C5FAS2 CRC64;

Query Match 61.5%; Score 123; DB 10; Length 343;
Best Local Similarity 64.7%; Pred. No. 9.9e-08;
Matches 23; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

Qy 1 GFLGLGMLGSGIVSNLLKMGHTVTVWNRATKCD 34
Dd 70 GFLGLGMLGSGIVSNLLKMGHTVTVWNRATKCD 103

RESULT 11
ID Q8LQJ7 PRELIMINARY; PRT; 343 AA.
AC Q8LQJ7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

```


DR Pfam; PF03446; NAD_binding_2; 1.
DR PRINTS; PR00076; 6PGDHRGNASE.
DR PROSITE; PS00895; 3 HYDROXYISOBUT_DH; 1.
KW Hypothetical protein.
SQ SEQUENCE 293 AA; 30496 MW; 2362B7AC4EB52038 CRC64;

Query Match 59.0%; Score 118; DB 10; Length 293;
Best Local Similarity 63.6%; Pred. No. 3.7e-07;
Matches 21; Conservative 4; Mismatches 8; Indels 0; Gaps 0

OY 1 GFGLGLMGSGIVSNLLKMGHTVTVMNRTAEKC 33
|||||::: : ||||: |||||
DB 4 GFGLGIMGKAMAANLLRHGRFVTVNRTLTKC 36
|||||::: : ||||: |||||

RESULT 13
Q97DX8 PRELIMINARY; PRT; 292 AA.
ID Q97DX8
AC Q97DX8
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dehydrogenase related to 3-hydroxyisobutyrate dehydrogenase, YKWC
DE B. subtilis ortholog.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL; AE007831; AAK81274.1; -
DR PIR; G97310; G97310.
DR GO; GO:0004616; F:phosphogluconate dehydrogenase (decarboxyla. . ; IEA.
DR GO; GO:0006098; P:pentose-phosphate shunt; IEA.
DR InterPro; IPR06183; 6PGD.
DR InterPro; IPR06115; 6PGD_NAD.
DR Pfam; PF03446; NAD_binding_2; 1.
DR PRINTS; PR00076; 6PGDHRGNASE.
KW Complete proteome.
SQ SEQUENCE 292 AA; 31542 MW; 7EFDCEC9D59EEB87 CRC64;

Query Match 54.5%; Score 109; DB 16; Length 292;
Best Local Similarity 62.5%; Pred. No. 5.5e-06;
Matches 20; Conservative 6; Mismatches 6; Indels 0; Gaps 0

OY 1 GFGLGLMGSGIVSNLLKMGHTVTVMNRTAEK 32
|||||::: : ||||: |||||
DB 9 GFIGTGMKGMIKNLLKGGYTIVRVNRTKEK 40
|||||::: : ||||: |||||

RESULT 14
Q97XZ7 PRELIMINARY; PRT; 289 AA.
ID Q97XZ7
AC Q97XZ7
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Oxidoreductase.
GV SSO1560.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI TaxID=2287;

Putative gamma hydroxybutyrate dehydrogenase.
GN Bl151A10.17.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoidae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RC Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1, BAC
RT clone.B151A10.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AP003413; BAB91904.1; -
DR Gramene; Q8LQJ7; -
DR GO; GO:0008442; F:3-hydroxyisobutyrate dehydrogenase activity; IEA.
DR GO; GO:0004616; F:phosphogluconate dehydrogenase (decarboxyla. . ; IEA.
DR GO; GO:0006098; P:pentose-phosphate shunt; IEA.
DR GO; GO:0006573; P:valine metabolism; IEA.
DR InterPro; IPR002204; 3hydroxisobut_dh.
DR InterPro; IPR006115; 6PGD_NAD.
DR Pfam; PF03446; NAD_binding_2; 1.
DR PROSITE; PS00895; 3 HYDROXYISOBUT_DH; 1.
SQ SEQUENCE 343 AA; 35481 MW; 195ABBEDA048F84D CRC64;

Query Match 60.5%; Score 121; DB 10; Length 343;
Best Local Similarity 64.7%; Pred. No. 1.8e-07;
Matches 22; Conservative 4; Mismatches 8; Indels 0; Gaps 0

OY 1 GFGLGLMGSGIVSNLLKMGHTVTVMNRTAEKCD 34
|||||::: : ||||: |||||
DB 54 GFGLGIMGAPWASNLINAGCDVTVMNRTSKCD 87
|||||::: : ||||: |||||

RESULT 12
Q84VC8 PRELIMINARY; PRT; 293 AA.
ID Q84VC8
AC Q84VC8
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Gamma hydroxybutyrate dehydrogenase-like protein (hypothetical
DE protein).
DE Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoidae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RC Cooper B., Hutchison D., Park S., Guimil S., Lugimbuhl P., Ellero C.,
RA Goff S., Glazebrook J.;
RA "Identification of Rice (Oryza sativa) Proteins Linked to the Cyclin-
RT Mediated Regulation of the Cell Cycle.";
RL Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Panicle;
RC Moriguchi K., Ito Y., Yamazaki Y., Kurata N.;
RT "Finding of various plant nuclear proteins using yeast nuclear
RT transportation trap system - a proteomal approach.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY224558; AAC72678.1; -
DR EMBL; AB110167; BAC78559.1; -
DR GO; GO:0008442; F:3-hydroxyisobutyrate dehydrogenase activity; IEA.
DR GO; GO:0004616; F:phosphogluconate dehydrogenase (decarboxyla. . ; IEA.
DR GO; GO:0006098; P:pentose-phosphate shunt; IEA.
DR GO; GO:0006573; P:valine metabolism; IEA.
DR InterPro; IPR002204; 3hydroxisobut_dh.
DR InterPro; IPR006183; 6PGD.
DR InterPro; IPR006115; 6PGD_NAD.


```

DR GO: GO:0004616; F:phosphoglucuronate dehydrogenase (decarboxyla. . .; IEA.
DR GO: GO:0006098; P:pentose-phosphate shunt; IEA.
DR InterPro: IPR006115; GPGD NAD.
DR Pfam: PF03446; NAD_binding_2; 1.
KW Complete proteome.
SQ SEQUENCE 155 AA; 16030 MW; 7FB9E77D3CC771A9 CRC64;

Query Match 52.5%; Score 105; DB 16; Length 155;
Best Local Similarity 59.4%; Pred. No. 9.3e-06;
Matches 19; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 3 LGLGLMSGIVSNLLKMGHTVTVMNRTAEKCD 34
   |||||: ||| ||||| ||||| : :
Db 7 LGLGAGTALARTLLKNGHVTVMNRTSRAE 38

RESULT 16
Q8T725
ID Q8T725 PRELIMINARY; PRT; 300 AA.
AC Q8T725;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE 3-hydroxyisobutyrate dehydrogenase.
GN MA0614.
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
```

RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=1929760; PubMed=11932238;
RA Galagan J.E., Nussbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Atncor D., Brown A.,
RA Allen N., Naylor J., Stange-Thomann N., DeArellano K., Johnson R.,
RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA Zimmer A., Barber P.D., Cann I., Graham D.P., Grasse D.A., Grasse A.M.

RA	Leigh J.A., Li W., Liu X., Mukhopadhyay B., Reeve J.N., Smith K.,
RA	Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA	Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA	Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA	Metcalf W.W., Birren B.
RT	"The genome of Methanosarcina acetivorans reveals extensive metabolic
RT	and physiological diversity."
RL	Genome Res. 12:532-542(2002).
RL	EMBL; AB3010722; AA004058.1 ; -.
DR	GO; GO:0004616; P:pentosephosphate dehydrogenase (decarboxyla. . ; IEA.
DR	GO; GO:0006098; P:phospho-phosphate shunt; IEA.
DR	InterPro; IPR006183; 6PGD.
DR	InterPro; IPR006115; 6PGD_NAD.
DR	Pfam; PF03446; NAD_binding_2; 1.
DR	PRINTS; PR00076; 6PGDHDRGNASE.
KW	Complete proteome.
KW	SEQUENCE 300 AA; 32287 MW; FB1B7ABAE8AERA7E CRC64;
Query Match	50.5%; Score 101; DB 17; Length 300;
Best Local Similarity	52.6%; Pred.No. 6.2e-05;
Matches 20; Conservative	5; Mismatches 13; Indels 0; Gaps 0;
QY	1 GFGLGIMGSGIVSNLLKMGHTVTVNVRTAEKCDFLIQ 38 : : : 13 GVIGLGIMSSFASNLRSRGYNHVYVNRTRKEKAQPLIE 50
Dd	

RESULT	17
Q88UE9	
ID	PRELIMINARY;
PRT;	286 AA.
AC	(T-EMBLrel. 24, Created)
DT	01-JUN-2003
DT	01-JUN-2003
DT	01-JUN-2003
DT	01-OCT-2003
DT	(T-EMBLrel. 24, Last sequence update)
DT	(T-EMBLrel. 25, Last annotation update)

3-hydroxyisobutyrate dehydrogenase (EC 1.1.1.31).

DE HIBD OR LP 2549.
 GN Lactobacillus plantarum.
 OS Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
 OC Lactobacillus.
 OX NCBI_TaxID=1590;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCIMB 8826 / WCFS1;
 RX MEDLINE=22480296; PubMed=12566566;
 RA Kluerebezem M., Boekhorst J., van Kranenburg R., Molenaar D.,
 RA Kuipers O.P., Leer R., Turchini R., Peters S.A., Sandbrink H.M.,
 RA Fiers M.W.E.J., Sniekema W., Klein Lankhorst R.M., Bron P.A.,
 RA Hoffer S.M., Nierop Groot M.N., Kerkhoven R., De Vries M., Ursing B.,
 RA De Vos W.M., Siezen R.J.;
 RL "Complete genome sequence of *Lactobacillus plantarum* WCFS1.";
 RT Proc. Natl. Acad. Sci. U.S.A. 100:1990-1995(2003).
 DR EMBL; AL935259; CAD64827.1; -;
 DR GO; GO:0008442; F:3-hydroxyisobutyrate dehydrogenase activity; IEA.
 DR GO; GO:0004420; F:hydroxymethylglutaryl-CoA reductase (NADPH) . . .; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0004616; F:phosphogluconate dehydrogenase (decarboxyla. . .; IEA.
 DR GO; GO:0009058; P:biosynthesis; IEA.
 DR GO; GO:0006098; P:pentose-phosphate shunt; IEA.
 DR InterPro; IPR006115; 6PGD NAD.
 DR InterPro; IPR002202; HMG-CoA red.
 DR Pfam; PF03446; NAD binding 2,1.
 DR PROSITE; PS00066; HMG COA REDUCTASE_1; 1.
 KW Oxidoreductase; Complete proteome.
 SQ SEQUENCE 286 AA; 30351 MW; AC9E968BDC222FE CRC64;

Query Match 50.0%; Score 100; DB 16; Length 286;
 Best Local Similarity 48.6%; Pred. No. 7.9e-05;
 Matches 18; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

Qy 1 GFGLGLMGSGIVSNLLKMGHTVTVMNRTAEKCDLFI 37
 Db 5 GWIGTGWGAIVRNLTAGYDVTVMNRTKSKADALV 41

RESULT 18

ID Q819E3 PRELIMINARY; PRT; 292 AA.
 AC Q819E3;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE 3-hydroxyisobutyrate dehydrogenase (EC 1.1.1.31).
 GN BC4042.
 OS Bacillus cereus (strain ATCC 14579 / DSM 31).
 OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=226900;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=22608415; PubMed=12721630;
 RA Ivanova N., Sorokin A., Anderson I., Galleron N., Candelson B.,
 RA Kapral V., Bhattacharyya A., Reznik G., Mikhailova N., Lapidus A.,
 RA Chu L., Mazur M., Goltzman E., Larsen N., D'Souza M., Walunas T.,
 RA Grechkin Y., Pusch G., Haselkorn R., Fonstein M., Ehrlich S.D.,
 RA Overbeek R., Kyrpides N.;
 RL "Genome sequence of *Bacillus cereus* and comparative analysis with
 RT *Bacillus anthracis*.";
 RL Nature 423:87-91(2003).
 DR EMBL; AE017011; AAP10961.1; -;
 DR GO; GO:0008442; F:3-hydroxyisobutyrate dehydrogenase activity; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0004616; F:phosphogluconate dehydrogenase (decarboxyla. . .; IEA.
 DR GO; GO:0006098; P:pentose-phosphate shunt; IEA.
 DR InterPro; IPR006183; 6PGD.
 DR InterPro; IPR006115; 6PGD NAD.
 DR Pfam; PF03446; NAD binding 2,1.
 DR PRINTS; PR00076; 6PGDHRGNASE.
 KW Oxidoreductase; Complete proteome.

SQ SEQUENCE 292 AA; 31987 MW; 8B04FC5E9A93423 CRC64;

Query Match 49.5%; Score 99; DB 16; Length 292;
 Best Local Similarity 47.4%; Pred. No. 0.00011;
 Matches 18; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

Qy 1 GFGLGLMGSGIVSNLLKMGHTVTVMNRTAEKCDLFIQ 38
 Db 9 GFGLGVMGKSNVYHLMQDGHKVVYVNRKAKTSLVQ 46

RESULT 19

ID Q811R8 PRELIMINARY; PRT; 291 AA.
 AC Q811R8;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Probable dehydrogenase.
 GN PA2199.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
 RA Brody L.I., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RL "Complete genome sequence of *Pseudomonas aeruginosa* PA01, an
 RT opportunistic pathogen.";
 RL Nature 406:959-964(2000).
 DR EMBL; AE004646; AAG05587.1; -;
 DR PIR; D83371; D83371.
 DR GO; GO:0008442; F:3-hydroxyisobutyrate dehydrogenase activity; IEA.
 DR GO; GO:0006573; P:valine metabolism; IEA.
 DR InterPro; IPR002204; 3hydroxisobut dh.
 DR PROSITE; PS00895; 3_HYDROXYISOBUT_DH; 1.
 KW Complete proteome.
 SQ SEQUENCE 291 AA; 30886 MW; F379B9582ACF81B4 CRC64;

Query Match 49.0%; Score 98; DB 16; Length 291;
 Best Local Similarity 52.9%; Pred. No. 0.00015;
 Matches 18; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

Qy 1 GFGLGLMGSGIVSNLLKMGHTVTVMNRTAEKCD 34
 Db 4 GFGLGGMGAAMATRLVQAGLEVTVMNRSAAACE 37

RESULT 20

ID Q88NR2 PRELIMINARY; PRT; 295 AA.
 AC Q88NR2;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE 3-hydroxyisobutyrate dehydrogenase family protein.
 GN PFI143.
 OS Pseudomonas putida (strain KT2440).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=160488;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=22423060; PubMed=12534463;
 RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
 RA Martins dos Santos V.A.P., Fouts D.E., Gill R.R., Pop M., Holmes M.,

```
RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,
RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,
RA Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazzes A.,
RA Utterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,
RA Lauber J., Sjepeandic D., Hohnsels J., Straetz M., Helm S.,
RA Kiewitz C., Eisen J., Timmis K.N., Duesterhoeft A., Tuemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440.";
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AE016778; AAN66768.1; -.
DR TIGR; PF1143; -.
DR GO; GO:0004616; F:phosphogluconate dehydrogenase (decarboxyla. .; IEA.
DR GO; GO:0006098; P:pentose-phosphate shunt; IEA.
DR InterPro; IPR006115; 6PGD NAD.
DR Pfam; PF03446; NAD_binding_2; 1.
DR Complete proteome.
KW SEQUENCE 295 AA; 30664 MW; 27E6E003982ED459 CRC64;
SQ
Query Match 48.0%; Score 96; DB 16; Length 295;
Best Local Similarity 51.5%; Pred. No. 0.00027;
Matches 17; Conservative 6; Mismatches 10; Indels 0; Gaps 0;
QY 1 GFLGLMGSGIVSNLLKMGHTVTWNRRTAEK 33
DB 9 GFAGTGLMGLEPCRRLLAAGYELTVWNRSPDK 41
RESULT 21
Q89HA0 PRELIMINARY; PRT; 313 AA.
AC Q89HA0;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Oxidoreductase.
GN BLR6094.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA 110;
RX Medline=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Ideawa K., Iiguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RL DNA Res. 9:189-197(2002).
DR EMBL; AP005957; BACS1359.1; -.
DR GO; GO:0004616; F:phosphogluconate dehydrogenase (decarboxyla. .; IEA.
DR GO; GO:0006098; P:pentose-phosphate shunt; IEA.
DR InterPro; IPR006183; 6PGD.
DR Pfam; PF03446; NAD_binding_2; 1.
DR PRINTS; PR00076; 6PGDHRGNASE.
DR Complete proteome.
KW SEQUENCE 313 AA; 32732 MW; 3B92612722105D51 CRC64;
SQ
Query Match 48.0%; Score 96; DB 16; Length 313;
Best Local Similarity 56.2%; Pred. No. 0.00029;
Matches 18; Conservative 6; Mismatches 8; Indels 0; Gaps 0;
QY 1 GFLGLMGSGIVSNLLKMGHTVTWNRRTAEK 32
DB 27 GYLGLMGTPMTRLLKAGYQSVWNRSEK 58
RESULT 22
Q81S35
```

```
Q81S35 PRELIMINARY; PRT; 317 AA.
AC Q81S35;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Dehydrogenase, putative.
GN BAI842.
OS Bacillus anthracis (strain Ames).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=198094;
RN [1]
RP SEQUENCE FROM N.A.
RX Medline=22608414; PubMed=12721629;
RA Read T.D., Peterson S.N., Tourasse N., Baillie L.W., Paulsen I.T.,
RA Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,
RA Holtzapple E.K., Okstad O.A., Helgason E., Ristone J., Wu M.,
RA Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.,
RA DeBoy R.T., Madupu R., Daugherty S.C., Durkin A.S., Haft D.H.,
RA Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,
RA Benton J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman J.F.,
RA Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman W.C.,
RA Hazen A., Cline R., Redmond C., Thwaite J.E., White O., Salzberg S.L.,
RA Thomason B., Friedlander A.M., Koehler T.M., Hanna P.C., Kolisto A.-B.,
RA Fraser C.M.;
RT "The genome sequence of Bacillus anthracis Ames and comparison to
RT closely related bacteria.";
RL Nature 423:81-86(2003).
DR EMBL; AE017029; AAP25747.1; -.
DR TIGR; BAI842; -.
DR GO; GO:0004616; F:phosphogluconate dehydrogenase (decarboxyla. .; IEA.
DR GO; GO:0006098; P:pentose-phosphate shunt; IEA.
DR InterPro; IPR006115; 6PGD NAD.
DR Pfam; PF03446; NAD_binding_2; 1.
DR Complete proteome.
KW SEQUENCE 317 AA; 33312 MW; CEA29415E4EF7A81 CRC64;
SQ
Query Match 48.0%; Score 96; DB 16; Length 317;
Best Local Similarity 51.4%; Pred. No. 0.00029;
Matches 18; Conservative 3; Mismatches 14; Indels 0; Gaps 0;
QY 3 LGLGLMGSGIVSNLLKMGHTVTWNRRTAEKCDLFI 37
DB 35 IGLGFMGQALAGAFMLNGHPTTLWNRRTAEKADYLI 69
RESULT 23
Q98GH7 PRELIMINARY; PRT; 475 AA.
AC Q98GH7;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 6-phosphogluconate dehydrogenase.
GN MLL3321.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX Medline=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Ideawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RL DNA Res. 7:331-338(2000).
DR EMBL; AP003001; BAB50239.1; -.
DR GO; GO:0004616; F:phosphogluconate dehydrogenase (decarboxyla. .; IEA.
DR GO; GO:0006098; P:pentose-phosphate shunt; IEA.
```

```
DR InterPro: IPR008927; 6DGDH_C_like.
DR InterPro: IPR006183; 6PGD.
DR InterPro: IPR006184; 6PGdom_BS.
DR InterPro: IPR006114; 6PGD_C.
DR InterPro: IPR006113; 6PGD decarbox.
DR InterPro: IPR006115; 6PGD_NAD.
DR Pfam: PF00393; 6PGD; 1.
DR Pfam: PF03446; NAD binding_2; 1.
DR PRINTS: PR00076; 6PGDHRGNASE.
DR TIGRfams: TIGR00873; gnd; 1.
DR PROSITE: PS00461; 6PGD; 1.
KW Complete proteome.
SQ SEQUENCE 475 AA; 50478 MW; DDCD80AE38E22DEE CRC64;

Query Match 48.0%; Score 96; DB 16; Length 475;
Best Local Similarity 42.1%; Pred. No. 0.00045;
Matches 16; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

QY 1 GFLGLMGSGIVSNLLKMGHTVTWNRTAEKCDLFIQ 38
   ||||| : : : : : ||| : ||| : ||| : ||| :
Db 7 GLIGLGTGMSNLALNIAEHGRIAVNRTAEVDAFVE 44

RESULT 24
Q82MW0 PRELIMINARY; PRT; 296 AA.
ID Q82MW0
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative dehydrogenase.
GN SAV1542.
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL: AP005027; BAC69253.1; -.
DR GO: GO:0004616; P:phosphogluconate dehydrogenase (decarboxyla...; IEA.
DR GO: GO:0006098; P:pentose-phosphate shunt; IEA.
DR InterPro: IPR006115; 6PGD_NAD.
DR Pfam: PF03446; NAD binding_2; 1.
KW Complete proteome.
SQ SEQUENCE 296 AA; 30626 MW; 4A235B663CE97764 CRC64;

Query Match 46.5%; Score 93; DB 16; Length 296;
Best Local Similarity 50.0%; Pred. No. 0.00057;
Matches 16; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 3 LGGLMGSGIVSNLLKMGHTVTWNRTAEKCD 34
   ||||| : : : : : ||||| : ||||| : ||||| :
Db 10 LGTGINGAMARNLNRAGHTVAVNRTAEKAE 41
```

```
RESULT 25
Q8FTI1 PRELIMINARY; PRT; 490 AA.
ID Q8FTI1
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative 6-phosphogluconate dehydrogenase.
GN CE1598.
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=152794;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
RA Kawarabayashi Y., Yamazaki J., Hino Y., Kikuchi H., Nakamura Y.,
RA Ikeo K., Suzuki M., Mashima J., Itoh T., Yamagishi A., Nishio Y.,
RA Usuda Y., Sugimoto S.;
RT "The entire genomic sequence of Corynebacterium efficiens YS-314.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF005219; BAC18398.1; -.
DR GO: GO:0004616; P:phosphogluconate dehydrogenase (decarboxyla...; IEA.
DR GO: GO:0006098; P:pentose-phosphate shunt; IEA.
DR InterPro: IPR008927; 6DGDH_C_like.
DR InterPro: IPR006183; 6PGD.
DR InterPro: IPR006184; 6PGdom_BS.
DR InterPro: IPR006114; 6PGD_C.
DR InterPro: IPR006113; 6PGD decarbox.
DR InterPro: IPR006115; 6PGD_NAD.
DR Pfam: PF00393; 6PGD; 1.
DR Pfam: PF03446; NAD binding_2; 1.
DR PRINTS: PR00076; 6PGDHRGNASE.
DR TIGRfams: TIGR00873; gnd; 1.
DR PROSITE: PS00461; 6PGD; 1.
KW Complete proteome.
SQ SEQUENCE 490 AA; 53121 MW; BA20D7E0CDD86761 CRC64;

Query Match 46.5%; Score 93; DB 16; Length 490;
Best Local Similarity 45.9%; Pred. No. 0.0011;
Matches 17; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

QY 1 GFLGLMGSGIVSNLLKMGHTVTWNRTAEKCDLFI 37
   ||||| : : : : : ||||| : ||||| : ||||| :
Db 17 GVUGLAVMGSLNLRNFARNGHVAVVNRSTDKTKLI 53

RESULT 26
O05188 PRELIMINARY; PRT; 43 AA.
ID O05188
AC O05188;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative 6-phospho-gluconate dehydrogenase (Fragment).
GN ORF2.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98015408; PubMed=9353924;
RA Mueller J., Schiel S., Ordal G.W., Saxild H.H.;
RT "Functional and genetic characterization of mcpC, which encodes a
RT third methyl accepting chemotaxis protein in Bacillus subtilis.";
RL Microbiology 143:3231-3240(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=93194799; PubMed=8449881;
RA Fajardo-Cavazos P., Salazar C., Nicholson W.L.;
RT "Molecular cloning and characterization of the Bacillus subtilis spore
```

```

RT photoproduct lyase (spl) gene, which is involved in repair of UV
RT radiation-induced DNA damage during spore germination.;
RL J. Bacteriol. 175:1735-1744(1993).
DR EMBL; X97385; CAA66053.1; -.
DR GO; GO:0008442; F:3-hydroxyisobutyrate dehydrogenase activity; IEA.
DR GO; GO:0004616; F:phosphogluconate dehydrogenase (decarboxyla. .; IEA.
DR GO; GO:0006098; P:pentose-phosphate shunt; IEA.
DR GO; GO:0006573; P:valine metabolism; IEA.
DR InterPro; IPR002204; 3hydroxisobut_dh.
DR IPR006115; 6PGD_NAD.
DR Pfam; PF03446; NAD_binding_2; 1.
DR PROSITE; PS00895; 3 HYDROXYISOBUT_DH; 1.
FT NON TER 43
FT 43
SQ SEQUENCE 43 AA; 4797 MW; 4E9F29DCB64945B CRC64;

Query Match 46.0%; Score 92; DB 2; Length 43;
Best Local Similarity 53.1%; Pred. No. 0.00012;
Matches 17; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 1 FLGLGLMGSGIVSNLLKMGHTVTVNRTAEK 32
Db 6 GFGLGVMGKSMASHILNDGHPVLVYTRTKK 37

RESULT 27
Q7WFF7 PRELIMINARY; PRT; 301 AA.
AC Q7WFF7;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative oxidoreductase.
GN BBA224.
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
RN [1]
RC STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640449; CAE34588.1; -.
KW Complete proteome.
SQ SEQUENCE 301 AA; 91AC4FBF73F6925F CRC64;

Query Match 46.0%; Score 92; DB 16; Length 301;
Best Local Similarity 58.1%; Pred. No. 0.00091;
Matches 18; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 2 FLGLGLMGSGIVSNLLKMGHTVTVNRTAEK 32
Db 19 FLGLGVMLPAGHARAGHQVTVNRTAAK 49

RESULT 28
Q7W489 PRELIMINARY; PRT; 302 AA.
AC Q7W489;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)

```

```

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative oxidoreductase.
GN BFP3779.
OS Bordetella parapertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=519;
RN [1]
RC STRAIN=12822 / ATCC BAA-587;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640434; CAE39062.1; -.
KW Complete proteome.
SQ SEQUENCE 302 AA; B9696B7BAD09CA9D CRC64;

Query Match 46.0%; Score 92; DB 16; Length 302;
Best Local Similarity 58.1%; Pred. No. 0.00092;
Matches 18; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 2 FLGLGLMGSGIVSNLLKMGHTVTVNRTAEK 32
Db 19 FLGLGVMLPAGHARAGHQVTVNRTAAK 49

RESULT 29
Q9SZE1 PRELIMINARY; PRT; 334 AA.
AC Q9SZE1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Putative dehydrogenase).
GN F19B15.150 OR A14G29120.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RC SEQUENCE FROM N.A.
RA Bevan M., Hilbert H., Braun M., Holzer E., Brandt A., Duisterhoef A.,
RA Bancroft I., Mewes H.W., Mayer K.F.X., Lemcke K., Schueller C.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RC SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RC SEQUENCE FROM N.A.
RA Hilbert H., Braun M., Holzer E., Brandt A., Duisterhoef A.,
RA Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RC SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RC SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Tang C.C., Toriumi M., Yu G.,

```

```

Query Match      46.0%; Score 92; DB 10; Length 334;
Best Local Similarity 45.9%; Pred.No. 0.001;
Matches 17; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY      1  GELGLGIMGSGIVSNLLKMGHTVTVMNRTAEKDLFI 37
      ||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB      41  GWIGTGWNGSMCGHLKAGIVTVFNRTISKAGTLI 77

RESULT 30
Q8PQK1 PRELIMINARY; PRT; 292 AA.
AC      Q8PQK1
DT      01-OCT-2002 (TrEMBLrel. 22, Created)
DT      01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT      01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE      Dehydrogenase.
GN      xAc0322
OS      Xanthomonas axonopodis (pv. citri).
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC      Xanthomonadaceae; Xanthomonas.
OC      NCBI_TaxID=92829;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=306 / ATCC 13902 / XV 101;
RX      MEDLINE=22022145; PubMed=12024217;
RA      ga Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA      Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA      Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA      Camarotte G., Cannavan F., Cardozo J., Chambergro F., Ciapina L.P.,
RA      Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA      Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA      Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA      Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA      Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA      Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2004, 07:28:53 ; Search time 59 Seconds
(without alignments)
2648.284 Million cell updates/sec

Title: US-10-067-482-2
Perfect score: 2866
Sequence: 1 MAVALSLRLGLVWGLGRYP.....AKALDQSDMSAVRYAIH 553

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

- Database : A_Geneseq_29Jan04.*
- 1: Geneseqp1980s.*
 - 2: Geneseqp1990s.*
 - 3: Geneseqp2000s.*
 - 4: Geneseqp2001s.*
 - 5: Geneseqp2002s.*
 - 6: Geneseqp2003as.*
 - 7: Geneseqp2003bs.*
 - 8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2858	99.7	553	6	Aa016415 Human nuc
2	2811	98.1	547	2	Aaw69240 Clone A07
3	2811	98.1	547	4	Aab90720 Human A07
4	2798	97.6	550	4	Aau21690 Novel hum
5	2798	97.6	550	7	Adc46331 Human neo
6	2396.5	83.6	472	6	Aae38240 Human enz
7	1404	49.0	269	4	Aau20515 Human sec
8	1404	49.0	269	4	Aau21813 Novel hum
9	1404	49.0	269	7	Adc46454 Human neo
10	1291	45.0	249	2	Aaw09404 Human hep
11	1291	45.0	249	6	Aau09982 Human hep
12	1291	45.0	260	4	Aau20553 Human sec
13	1291	45.0	260	4	Aau21747 Novel hum
14	1286	44.9	250	7	Adc46388 Human neo
15	1286	44.9	250	7	Aau21590 Novel hum
16	1286	44.9	250	7	Adc46231 Human neo
17	769.5	26.8	574	4	Aab60930 Drosophil
18	733	25.6	360	5	Aae19927 Corn 3-hy
19	720	25.1	290	6	Aae36391 Tulip pis
20	695	24.2	289	6	Aae36392 A. thalia
21	682	23.8	345	5	Aae19929 Soybean 3
22	679	23.7	289	3	Aab19491 Amino aci
23	667.5	23.3	353	3	Aggl0423 Arabidops
24	667	23.3	343	3	Aggl0424 Arabidops
25	629	21.9	285	3	Aggl0425 Arabidops

26	537.5	18.8	247	5	AAE19926	Aae19926 Corn 3-hy
27	399	13.9	168	3	AAG41222	Aag41222 Arabidops
28	384	13.4	312	3	AAG20956	Aag20956 Arabidops
29	384	13.4	334	3	AAG20955	Aag20955 Arabidops
30	384	13.4	336	3	AAG20954	Aag20954 Arabidops
31	375	13.1	292	6	ABU24159	Abu24159 Protein e
32	367	12.8	299	3	AAE15899	Aae15899 E. coli p
33	361	12.6	158	3	AAG41223	Aag41223 Arabidops
34	352	12.3	154	3	AAG41224	Aag41224 Arabidops
35	350	12.2	296	6	ABU28970	Abu28970 Protein e
36	343.5	12.0	299	6	ADA36120	Ada36120 Acinetoba
37	342	11.9	289	3	AAE15899	Aae15899 E. coli p
38	342	11.9	289	3	AAE15899	Aae15899 E. coli p
39	342	11.9	289	3	AAE15899	Aae15899 E. coli p
40	340	11.9	289	3	AAE15899	Aae15899 E. coli p
41	340	11.9	289	3	AAE15899	Aae15899 E. coli p
42	334.5	11.7	295	7	ADG95482	Adg95482 E. faeciu
43	331	11.5	286	5	ADG95482	Adg95482 E. faeciu
44	331	11.5	286	5	ADG95482	Adg95482 E. faeciu
45	330	11.5	286	5	ADG95482	Adg95482 E. faeciu
46	330	11.5	286	5	ADG95482	Adg95482 E. faeciu
47	329	11.5	289	3	AAE15899	Aae15899 E. coli p
48	329	11.5	289	3	AAE15899	Aae15899 E. coli p
49	328	11.4	347	3	AAE15899	Aae15899 E. coli p
50	327.5	11.4	290	6	ABU17066	Abu17066 Protein e
51	327.5	11.4	347	3	AAE15899	Aae15899 E. coli p
52	324.5	11.3	288	6	ABU38377	Abu38377 Protein e
53	318	11.1	296	6	ABU38377	Abu38377 Protein e
54	316	11.0	293	5	ABU38377	Abu38377 Protein e
55	314.5	11.0	277	6	ABU38377	Abu38377 Protein e
56	314	11.0	297	6	ABU38377	Abu38377 Protein e
57	313.5	10.9	161	7	ABU38377	Abu38377 Protein e
58	313.5	10.9	301	6	ABU38377	Abu38377 Protein e
59	301.5	10.5	292	4	AAE15899	Aae15899 E. coli p
60	301.5	10.5	292	4	AAE15899	Aae15899 E. coli p
61	301.5	10.5	297	5	ABU14750	Abu14750 Protein e
62	296.5	10.3	300	3	AAE15899	Aae15899 E. coli p
63	296.5	10.3	300	3	AAE15899	Aae15899 E. coli p
64	295	10.3	299	5	ABU14750	Abu14750 Protein e
65	291.5	10.2	292	6	ABU14750	Abu14750 Protein e
66	290.5	10.1	298	6	ABU33528	Abu33528 Protein e
67	289.5	10.1	296	3	AAE15899	Aae15899 E. coli p
68	289.5	10.1	296	3	AAE15899	Aae15899 E. coli p
69	287.5	10.0	303	6	ADA33592	Ada33592 Acinetoba
70	285	9.9	298	6	ADA33592	Ada33592 Acinetoba
71	281	9.8	291	6	ABU45556	Abu45556 Protein e
72	281	9.8	336	5	ABU45556	Abu45556 Protein e
73	281	9.8	336	5	ABU45556	Abu45556 Protein e
74	281	9.8	336	5	ABU45556	Abu45556 Protein e
75	279.5	9.8	292	6	ABU45556	Abu45556 Protein e
76	279.5	9.8	302	6	ABU45556	Abu45556 Protein e
77	278.5	9.7	324	4	ABU70141	Abu70141 Drosophil
78	277	9.7	302	6	ABU21583	Abu21583 Protein e
79	272.5	9.5	292	4	ABU52479	Abu52479 Escherich
80	272	9.5	292	4	ABU52479	Abu52479 Escherich
81	269	9.4	287	6	ABU17331	Abu17331 Protein e
82	266	9.3	326	6	ABU17331	Abu17331 Protein e
83	258.5	9.0	342	5	AAE19928	Aae19928 Rice 3-hy
84	250.5	8.7	158	3	AAE19928	Aae19928 Rice 3-hy
85	250.5	8.7	180	3	AAE19928	Aae19928 Rice 3-hy
86	250.5	8.7	182	3	AAE19928	Aae19928 Rice 3-hy
87	248.5	8.7	287	5	AAE19928	Aae19928 Rice 3-hy
88	248.5	8.7	287	5	AAE19928	Aae19928 Rice 3-hy
89	235	8.4	350	5	AAE19928	Aae19928 Rice 3-hy
90	234.5	8.2	294	7	ADB80011	Adb80011 Mycobacte
91	214.5	7.5	551	4	ABG29536	Abg29536 Novel hum
92	209	7.3	426	4	ABG17925	Abg17925 Novel hum
93	199.5	7.0	305	6	ABG17925	Abg17925 Novel hum
94	193	6.7	475	4	ABG17925	Abg17925 Novel hum
95	190.5	6.6	202	5	ABG17925	Abg17925 Novel hum
96	189	6.6	202	5	ABG17925	Abg17925 Novel hum
97	189	6.6	203	3	AAE19928	Aae19928 Rice 3-hy
98	189	6.6	203	3	AAE19928	Aae19928 Rice 3-hy

99	189	6.6	203	4	AAB70073	Human sec
100	189	6.6	203	4	AAB92800	Human pro
ALIGNMENTS						
RESULT 1						
ID	AAO16415	standard; protein; 553 AA.				
XX	AAO16415	(first entry)				
AC	AAO16415	Human nucleic acid-associated protein (NAAP) - SEQ ID No 12.				
DT	10-APR-2003	(first entry)				
DE	Human nucleic acid-associated protein (NAAP) - SEQ ID No 12.					
XX	Human; nucleic acid-associated protein; NAAP; arteriosclerosis;					
KW	cell proliferative disorder; atherosclerosis; cirrhosis; hepatitis; AIDS;					
KW	cancer; developmental disorder; renal tubular acidosis; anaemia; asthma;					
KW	mental retardation; neurological disorder; Alzheimer's disease; epilepsy;					
KW	Parkinson's disease; autoimmune disorder; inflammatory disorder; allergy;					
KW	Crohn's disease; transgenic animal; animal model.					
XX	Homo sapiens.					
OS	WO2003000864-A2.					
PN	03-JAN-2003.					
PD	20-JUN-2002; 2002WO-US021179.					
XX	22-JUN-2001; 2001US-0300518P.					
PR	29-JUN-2001; 2001US-0301787P.					
PR	29-JUN-2001; 2001US-0301792P.					
PR	29-JUN-2001; 2001US-0301893P.					
PR	29-JUN-2001; 2001US-0301893P.					
PR	06-JUL-2001; 2001US-0303405P.					
PR	06-JUL-2001; 2001US-0303442P.					
PR	15-MAR-2002; 2002US-0364438P.					
XX	(INCY-) INCYTE GENOMICS INC.					
PA	Gandhi AR, Swarnakar A, Hafalia AJA, Warren BA, Emerling BM;					
XX	Arvizu CS, Ison CH, Honchell CB, Lee EA, Yue H, Forsythe IJ;					
PI	Ramkumar J, Griffin JA, Yang J, Sanjanwala MM, Baughn MK;					
PI	Borowsky ML, Yao MG, Walia NK, Bandman O, Lal PG, Becha SD, Lee SY;					
PI	Richardson TW, Elliott VS, Luo W, Tang YT, Zebardjian Y, Lu Y;					
XX	WPT; 2003-201420/19.					
DR	N-PSDB; AAL51565.					
XX	New nucleic acid-associated proteins and polynucleotides, useful for					
PT	diagnosing, treating or preventing cell proliferative (e.g. cancer),					
PT	neurological (e.g. epilepsy or Parkinson's disease), or autoimmune					
PT	disorders (e.g. AIDS).					
XX	Claim 1; Page 227-228; 312pp; English.					
XX	The invention comprises the amino acid and coding sequences of human					
CC	nucleic acid-associated proteins (NAAP). The DNA and protein sequences of					
CC	the invention are useful for diagnosing, treating or preventing disorders					
CC	associated with aberrant expression of NAAP, such as: cell proliferative					
CC	disorders (e.g. arteriosclerosis, atherosclerosis, cirrhosis, hepatitis					
CC	or cancer); developmental disorders (e.g. renal tubular acidosis, anaemia					
CC	or mental retardation); neurological disorders (e.g. Alzheimer's disease,					
CC	Parkinson's disease or epilepsy); and autoimmune/inflammatory disorders					
CC	(e.g. AIDS, allergies, asthma or Crohn's disease). The DNA sequences of					
CC	the invention are useful for creating transgenic animals to model human					
CC	disease. The present amino acid sequence represents a human nucleic acid-					
CC	associated protein of the invention					
XX	Sequence 553 AA;					

Query Match		99.7%;	Score 2858;	DB 6;	Length 553;
Best Local Similarity		99.8%;	Pred. No. 2.7e-254;		
Matches 552;		Conservative	0;	Mismatches 1;	Indels 0; Gaps 0;
QY	1	MAAVSLRLGDLVWGKLGRIYPPWPGKIVNPPKDLKKPRGKKCFVFFFGTDEHAWIKVEOL	60		
DB	1	MAAVSLRLGDLVWGKLGRIYPPWPGKIVNPPKDLKKPRGKKCFVFFFGTDEHAWIKVEOL	60		
QY	61	KPYHAHKEMIKINIKGKRFQQAQVDAVEEFLRAKGDQTSNHNSSDDKNRRNSSERSRP	120		
DB	61	KPYHAHKEMIKINIKGKRFQQAQVDAVEEFLRAKGDQTSNHNSSDDKNRRNSSERSRP	120		
QY	121	NSGDEKRLKLSLEGKVKKNMGEGKRVSSGSGSERGSKPLKRAQEQSPKRGPPPKDEKD	180		
DB	121	NSGDEKRLKLSLEGKVKKNMGEGKRVSSGSGSERGSKPLKRAQEQSPKRGPPPKDEKD	180		
QY	181	LTIPSSSTVKGMMAGMAGMAAFKWOPTASEPVKQADPHFHFLLSQTEKPAVCYQAITKKLK	240		
DB	181	LTIPSSSTVKGMMAGMAGMAAFKWOPTASEPVKQADPHFHFLLSQTEKPAVCYQAITKKLK	240		
QY	241	ICEEETGTSIOAADSTAVNGSITPTDKKIGFLGLGMSGGIVSNLLKMGHTVTVNRTA	300		
DB	241	ICEEETGTSIOAADSTAVNGSITPTDKKIGFLGLGMSGGIVSNLLKMGHTVTVNRTA	300		
QY	301	EKCDLFIQEGARLGRTPAEVWSTCDITFACVSDPKAAKDLVLGPGSVLQIRPGKCYVDM	360		
DB	301	EKCDLFIQEGARLGRTPAEVWSTCDITFACVSDPKAAKDLVLGPGSVLQIRPGKCYVDM	360		
QY	361	STVDADTVTELAQVIIVSRGRFLEAPVSGNQQLSNDGMLVILAAGDRGLYEDCSSCFQAM	420		
DB	361	STVDADTVTELAQVIIVSRGRFLEAPVSGNQQLSNDGMLVILAAGDRGLYEDCSSCFQAM	420		
QY	421	GKTSFFLGEVGNAAKMLIVNMVQSGFMATIAEGLTLAHTVGTQSQQTLLDLINQGQLASI	480		
DB	421	GKTSFFLGEVGNAAKMLIVNMVQSGFMATIAEGLTLAHTVGTQSQQTLLDLINQGQLASI	480		
QY	481	FLDQKQCNILQGNFKPDFYLYIQKDLRLAIALGDVNHPTPMAAAANEVYKRAKALDOS	540		
DB	481	FLDQKQCNILQGNFKPDFYLYIQKDLRLAIALGDVNHPTPMAAAANEVYKRAKALDOS	540		
QY	541	DNDMSAVRAYIHH 553			
DB	541	DNDMSAVRAYIHH 553			
RESULT 2					
AAW69240					
ID	AAW69240 standard; protein; 547 AA.				
XX	AAW69240;				
DT	21-OCT-1998 (first entry)				
XX	Clone AQ73_3 protein sequence.				
DE	Secreted protein; nutritional source; cell proliferation activity;				
KW	cell differentiation activity; immune stimulant; tissue growth activator;				
KW	haematopoiesis regulator; anti-inflammatory; tumour invasion suppressor;				
KW	tumour inhibitor; clone AQ73_3.				
OS	Homo sapiens.				
XX	WO9825962-A2.				
PN	18-JUN-1998.				
PD					
XX	12-DEC-1997; 97WO-US023224.				
PF					
XX	13-DEC-1996; 96US-00766263.				
XX	11-DEC-1997; 97US-00989232.				
XX	(GEMY) GENETICS INST INC.				
PA					

XX PI Jacobs K, McCoy JM, Lavallie ER, Racie LA, Merberg D, Treacy M;
 XX PI Spaulding V, Agostino MJ;
 XX DR WPI; 1998-362424/31.
 XX DR N-PSDB; AAV44859.
 XX PT New isolated polynucleotides - obtained from human adult testis, human
 XX PT adult ovary, human adult brain and human adult heart cDNA libraries.
 XX XX Claim 15; Page 66-67; 108pp; English.
 XX CC This sequence is secreted protein, encoded by a polynucleotide of the
 CC invention. The DNA was isolated from a human adult ovary cDNA library,
 CC and is designated clone AQ73_3. The DNA sequences and encoded
 CC polypeptides can be used as nutritional sources or supplements, or may
 CC exhibit e.g. cytokine and cell proliferation/differentiation activity,
 CC immune stimulating or suppressing activity, haematopoiesis regulating
 CC activity, receptor/ligand activity, anti-inflammatory activity,
 CC activin/inhibin activity, chemostatic/chemokinetic activity,
 CC cadherin/tumour invasion suppressor activity, tissue growth activity,
 CC tumour inhibition activity or other activities
 XX SQ Sequence 547 AA;
 Query Match 98.1%; Score 2811; DB 2; Length 547;
 Best Local Similarity 98.7%; Pred. No. 5.8e-250;
 Matches 546; Conservative 0; Mismatches 1; Indels 6; Gaps 1;
 QY 1 MAASVRLGDLVWGKLGKRYPPGKLVNPPKDLKPRGKKCFVFKFGTEHAWIKVEQL 60
 Db 1 MAASVRLGDLVWGKLGKRYPPGKLVNPPKDLKPRGKKCFVFKFGTEHAWIKVEQL 60
 QY 61 KPYHAHEMIKINKGRFQQAQVDAVEEFLRAKGDQTSNHSDDKNNRNSSEERSRP 120
 Db 61 KPYHAHEMIKINKGRFQQAQVDAVEEFLRAKGDQTSNHSDDKNNRNSSEERSRP 120
 QY 121 NSGDEKRLSLSEGKVKKNMGEKKRVSSGSSGSKPLKRAQEQSPRKRPPKDEKD 180
 Db 121 NSGDEKRLSLSEGKVKKNMGEKKRVSSGSSGSKPLKRAQEQSPRKRPPKDEKD 180
 QY 181 LTIPESTVKGMMAGPMAAFKQWPTTASEPVKDDPHHFLLSQTEPAVCYQAITKKL 240
 Db 181 LTIPESTVKGMMAGPMAAFKQWPTTASEPVKDDPHHFLLSQTEPAVCYQAITKKL 240
 QY 241 ICEETGTSIQADSTAVNGSIITDCKIGLGLGSGIVSNLLKMGHTVTVNRTA 300
 Db 241 ICEETGTSIQADSTAVNGSIITDCKIGLGLGSGIVSNLLKMGHTVTVNRTA 300
 QY 301 EKCDLFIQEGARLGRTPAEVSTCDITFACVSDPKAAKDLVLGPGSVLQGIKPCYVDM 360
 Db 301 EK-----EGARLGRTPAEVSTCDITFACVSDPKAAKDLVLGPGSVLQGIKPCYVDM 360
 QY 361 STVDADITVTELAQVTSRGGFLRPAVSGNQLSNDGMLVILAGDRLGEDSCSCFOAM 420
 Db 361 STVDADITVTELAQVTSRGGFLRPAVSGNQLSNDGMLVILAGDRLGEDSCSCFOAM 420
 QY 421 KTSFFLGEVGNAAKMLIVNMVQGSFMATIAEGLTLAHVTGQSQOQLDILNQGOLASI 480
 Db 421 KTSFFLGEVGNAAKMLIVNMVQGSFMATIAEGLTLAHVTGQSQOQLDILNQGOLASI 480
 QY 481 FLDOCKQNILQGNPKDPFLKYI QKDLRLAIALGDVANNHPTPMAAANEVYKRAKALDQS 540
 Db 481 FLDOCKQNILQGNPKDPFLKYI QKDLRLAIALGDVANNHPTPMAAANEVYKRAKALDQS 540
 QY 541 DNDMSAVYAYIH 553
 Db 541 DNDMSAVYAYIH 553
 QY 535 DNDMSAVYAYIH 547
 Db 535 DNDMSAVYAYIH 547

RESULT 3
 AAB90720
 ID AAB90720 standard; protein; 547 AA.

XX AC AAB90720;
 XX DT 07-JUN-2001 (first entry)
 XX DE Human AQ73_3 protein sequence SEQ ID 137.
 XX KW Human; secreted protein; nutrient; cytokine modulator; proliferation;
 KW differentiation; immune system modulator; tissue growth; chemotactic;
 KW haemostatic; thrombolytic; anti-inflammatory; tumour inhibition;
 KW haematopoiesis.
 XX OS Homo sapiens.
 XX PN WO200119988-A1.
 XX XX 22-MAR-2001.
 XX PF 14-SEP-2000; 2000WO-US025135.
 XX PR 17-SEP-1999; 99US-00398829.
 XX XX (GEMY) GENETICS INST INC.
 XX PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
 PI Merberg D, Treacy M, Bowman MK, Spaulding V, Agostino MJ;
 XX DR WPI; 2001-244801/25.
 XX DR N-PSDB; AAF98456.
 XX PT Isolated nucleic acids encoding polypeptides, useful for modulating e.g.
 PT cytokine and cell proliferation/differentiation activity, the immune
 PT system and hematopoiesis regulating activity.
 XX PS Disclosure; Page 468-470; 557pp; English.
 XX CC Human cDNA clones represented in AAF98374 - AAF98489 encode secreted
 CC proteins AAB90667 - AAB90750. The cDNA clones are isolated from various
 CC tissue types, and may be used in the prevention, treatment and diagnosis
 CC of diseases associated with inappropriate protein expression. The
 CC polypeptides and nucleic acids may be used as nutrients or to modulate
 CC cytokine and cell proliferation/differentiation activity and may also be
 CC involved in modulation of the immune system. The cDNA sequences,
 CC proteins, their agonists and/or antagonists exhibit haematopoiesis
 CC regulating activity; tissue growth activity; activin/inhibin activity;
 CC chemotactic/chemokinetic activity; haemostatic and thrombolytic activity;
 CC receptor/ligand activity; anti-inflammatory activity; haematopoiesis
 CC activity; cadherin/tumour suppressor activity; and/or tumour inhibition
 CC activity. Included in the invention are probes represented in AAF98490 -
 CC AAF98572 which are specific for the cDNA clones encoding the secreted
 CC proteins
 XX SQ Sequence 547 AA;
 Query Match 98.1%; Score 2811; DB 4; Length 547;
 Best Local Similarity 98.7%; Pred. No. 5.8e-250;
 Matches 546; Conservative 0; Mismatches 1; Indels 6; Gaps 1;
 QY 1 MAASVRLGDLVWGKLGKRYPPGKLVNPPKDLKPRGKKCFVFKFGTEHAWIKVEQL 60
 Db 1 MAASVRLGDLVWGKLGKRYPPGKLVNPPKDLKPRGKKCFVFKFGTEHAWIKVEQL 60
 QY 61 KPYHAHEMIKINKGRFQQAQVDAVEEFLRAKGDQTSNHSDDKNNRNSSEERSRP 120
 Db 61 KPYHAHEMIKINKGRFQQAQVDAVEEFLRAKGDQTSNHSDDKNNRNSSEERSRP 120
 QY 121 NSGDEKRLSLSEGKVKKNMGEKKRVSSGSSGSKPLKRAQEQSPRKRPPKDEKD 180
 Db 121 NSGDEKRLSLSEGKVKKNMGEKKRVSSGSSGSKPLKRAQEQSPRKRPPKDEKD 180
 QY 181 LTIPESTVKGMMAGPMAAFKQWPTTASEPVKDDPHHFLLSQTEPAVCYQAITKKL 240
 Db 181 LTIPESTVKGMMAGPMAAFKQWPTTASEPVKDDPHHFLLSQTEPAVCYQAITKKL 240

Thu Sep 16 09:36:51 2004

us-10-067-482-2.rag

QY	241	ICEETGTSIQADSTAVNGSITPTDKKIGFLGLGMSGIVSNLLKMGHTVTVWNRTA	300
Db	241	ICEETGTSIQADSTAVNGSITPTDKKIGFLGLGMSGIVSNLLKMGHTVTVWNRTA	300
QY	301	EKCDLFIQEGARLGTAEVNSTCDITFACVSDPKAAKDLVLGPGSVLQGIKIRPGKCYVDM	360
Db	301	EK-----EGARLGTAEVNSTCDITFACVSDPKAAKDLVLGPGSVLQGIKIRPGKCYVDM	354
QY	361	STVDADTVTELAQIVTSRGRLEAPVSGNQQLSNDGMLVILAAGDRGLYEDCSSCFQAM	420
Db	355	STVDADTVTELAQIVTSRGRLEAPVSGNQQLSNDGMLVILAAGDRGLYEDCSSCFQAM	414
QY	421	GKTSFPLGEVGNAAKMWLVNVMVQGSFMATIAEGLTLAHVTGQSQTLLDILNQGLAST	480
Db	415	GKTSFPLGEVGNAAKMWLVNVMVQGSFMATIAEGLTLAQTGQSQTLLDILNQGLAST	474
QY	481	FLDQKQNTILQGNFKPDFLYIKYIOKDLRLAIALGDVNHPTMAAANAEVYKRAKALDOS	540
Db	475	FLDQKQNTILQGNFKPDFLYIKYIOKDLRLAIALGDVNHPTMAAANAEVYKRAKALDOS	534
QY	541	DNDMSAVYRAYIH	553
Db	535	DNDMSAVYRAYIH	547
RESULT 4			
AAU21690	AAU21690 standard; protein; 550 AA.		
XX	AC	AAU21690;	
XX	DT		
XX	DD	04-DEC-2001 (first entry)	
DE	XX	Novel human neoplastic disease associated polypeptide #123.	
XX	XX	Human; neoplastic disease associated polypeptide; cancer;	
KW	KW	hyperproliferative disorder; neural disorder; immune system disorder;	
KW	KW	muscular disorder; reproductive disorder; gastrointestinal disorder;	
KW	KW	pulmonary disorder; cardiovascular disorder; renal disorder;	
KW	KW	neuroprotective; cytostatic; anti inflammatory; vasotropic.	
XX	OS	Homo sapiens.	
XX	PN	WO200155163-A1.	
XX	PD	02-AUG-2001.	
XX	PF	17-JAN-2001; 2001WO-US001358.	
XX	PR	31-JAN-2000; 2000US-0179065P.	
PR	PR	04-FEB-2000; 2000US-0180628P.	
PR	PR	24-FEB-2000; 2000US-0184664P.	
PR	PR	02-MAR-2000; 2000US-0186350P.	
PR	PR	16-MAR-2000; 2000US-0189874P.	
PR	PR	17-MAR-2000; 2000US-0190076P.	
PR	PR	18-APR-2000; 2000US-01981123P.	
PR	PR	19-MAY-2000; 2000US-0205515P.	
PR	PR	07-JUN-2000; 2000US-0209467P.	
PR	PR	28-JUN-2000; 2000US-0214886P.	
PR	PR	30-JUN-2000; 2000US-0215135P.	
PR	PR	07-JUL-2000; 2000US-0216664P.	
PR	PR	07-JUL-2000; 2000US-0216880P.	
PR	PR	11-JUL-2000; 2000US-0217487P.	
PR	PR	11-JUL-2000; 2000US-0217496P.	
PR	PR	14-JUL-2000; 2000US-0218290P.	
PR	PR	26-JUL-2000; 2000US-0220963P.	
PR	PR	26-JUL-2000; 2000US-0220964P.	
PR	PR	14-AUG-2000; 2000US-0224518P.	
PR	PR	14-AUG-2000; 2000US-0224519P.	
PR	PR	14-AUG-2000; 2000US-0225213P.	
PR	PR	14-AUG-2000; 2000US-0225214P.	
PR	PR	14-AUG-2000; 2000US-0225266P.	

PR	14-AUG-2000;	2000US-0225267P.
PR	14-AUG-2000;	2000US-0225268P.
PR	14-AUG-2000;	2000US-0225270P.
PR	14-AUG-2000;	2000US-0225447P.
PR	14-AUG-2000;	2000US-0225577P.
PR	14-AUG-2000;	2000US-0225758P.
PR	14-AUG-2000;	2000US-0225759P.
PR	18-AUG-2000;	2000US-0226279P.
PR	22-AUG-2000;	2000US-0226681P.
PR	22-AUG-2000;	2000US-0226868P.
PR	22-AUG-2000;	2000US-0227182P.
PR	23-AUG-2000;	2000US-0227009P.
PR	30-AUG-2000;	2000US-0228924P.
PR	01-SEP-2000;	2000US-0229287P.
PR	01-SEP-2000;	2000US-0229343P.
PR	01-SEP-2000;	2000US-0229344P.
PR	01-SEP-2000;	2000US-0229345P.
PR	05-SEP-2000;	2000US-0229509P.
PR	05-SEP-2000;	2000US-0229513P.
PR	06-SEP-2000;	2000US-0230437P.
PR	06-SEP-2000;	2000US-0230438P.
PR	08-SEP-2000;	2000US-0231242P.
PR	08-SEP-2000;	2000US-0231243P.
PR	08-SEP-2000;	2000US-0231244P.
PR	08-SEP-2000;	2000US-0231413P.
PR	08-SEP-2000;	2000US-0231414P.
PR	08-SEP-2000;	2000US-0232080P.
PR	08-SEP-2000;	2000US-0232081P.
PR	12-SEP-2000;	2000US-0231968P.
PR	14-SEP-2000;	2000US-0232397P.
PR	14-SEP-2000;	2000US-0232398P.
PR	14-SEP-2000;	2000US-0232399P.
PR	14-SEP-2000;	2000US-0232400P.
PR	14-SEP-2000;	2000US-0232401P.
PR	14-SEP-2000;	2000US-0233063P.
PR	14-SEP-2000;	2000US-0233064P.
PR	14-SEP-2000;	2000US-0233065P.
PR	21-SEP-2000;	2000US-0234223P.
PR	21-SEP-2000;	2000US-0234274P.
PR	25-SEP-2000;	2000US-0234997P.
PR	25-SEP-2000;	2000US-0234998P.
PR	26-SEP-2000;	2000US-0235484P.
PR	27-SEP-2000;	2000US-0235834P.
PR	27-SEP-2000;	2000US-0235836P.
PR	29-SEP-2000;	2000US-0236327P.
PR	29-SEP-2000;	2000US-0236367P.
PR	29-SEP-2000;	2000US-0236368P.
PR	29-SEP-2000;	2000US-0236369P.
PR	29-SEP-2000;	2000US-0236370P.
PR	02-OCT-2000;	2000US-0236802P.
PR	02-OCT-2000;	2000US-0237037P.
PR	02-OCT-2000;	2000US-0237038P.
PR	02-OCT-2000;	2000US-0237039P.
PR	02-OCT-2000;	2000US-0237040P.
PR	13-OCT-2000;	2000US-0239935P.
PR	13-OCT-2000;	2000US-0239937P.
PR	20-OCT-2000;	2000US-0240960P.
PR	20-OCT-2000;	2000US-0241221P.
PR	20-OCT-2000;	2000US-0241785P.
PR	20-OCT-2000;	2000US-0241786P.
PR	20-OCT-2000;	2000US-0241787P.
PR	20-OCT-2000;	2000US-0241808P.
PR	20-OCT-2000;	2000US-0241809P.
PR	01-NOV-2000;	2000US-0241826P.
PR	08-NOV-2000;	2000US-0246474P.
PR	08-NOV-2000;	2000US-0246475P.
PR	08-NOV-2000;	2000US-0246476P.
PR	08-NOV-2000;	2000US-0246477P.
PR	08-NOV-2000;	2000US-0246478P.
PR	08-NOV-2000;	2000US-0246523P.
PR	08-NOV-2000;	2000US-0246524P.
PR	08-NOV-2000;	2000US-0246525P.

PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246612P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249277P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 01-DEC-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 05-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 06-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 11-DEC-2000; 2000US-0251990P.
 PR 05-JAN-2001; 2001US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 (HUMA-) HUMAN GENOME SCI INC.
 Rosen CA, Barash SC, Ruben SM;
 WPI; 2001-465558/50.
 N-PSDB; AAS34889.
 Novel polypeptides and polynucleotides useful as diagnostic reagents to
 diagnose diseases or disorders associated with aberrant expression or
 activity of polypeptides, and for treating cancers, rheumatoid arthritis.
 Claim 11; SEQ ID NO 417; 687pp; English.
 The present invention relates to the isolation of novel human neoplastic
 disease associated polypeptides, and cDNA (AAS34767-AAS35050) and DNA
 sequences encoding for these polypeptides. The sequences of the invention
 are useful in the diagnosis, treatment, prevention and/or prognosis of
 disorders involving neoplastic disease such as hyperproliferative
 disorders (e.g. leukaemia, bone cancer, bladder cancer, brain stem
 glioma, adult liver cancer, childhood cerebellar astrocytoma, or
 Hodgkin's lymphoma). The sequences of the invention may also be useful
 for treating other disorders such as neural disorders, immune system
 disorders, pulmonary disorders, reproductive disorders, gastrointestinal
 disorders, muscular disorders, cardiovascular disorders and renal
 disorders. The polynucleotide sequences of the invention are also useful
 in gene therapy. AAU1568-AAU1851 represent the novel human neoplastic
 disease associated polypeptides of the invention. Note: The sequence data
 for this patent did not form part of the printed specification, but was
 obtained in electronic format directly from WIPO at
 ftp.wipo.int/pub/published_pct_sequences
 Sequence 550 AA;

Query Match 97.6%; Score 2798; DB 4; Length 550;
 Best Local Similarity 98.7%; Pred. No. 9.3e-249;
 Matches 543; Conservative 0; Mismatches 1; Indels 6; Gaps 1;
 QY 4 VSLRLGDLVWGKLGRIYPPWPGKIVNPPKDLKKPKGKCKFFVKFGTGHAIKVEQLKPY 63
 DB 7 VSLRLGDLVWGKLGRIYPPWPGKIVNPPKDLKKPKGKCKFFVKFGTGHAIKVEQLKPY 66
 QY 64 HAHKEEMIKINKGRFOQAVDAVEEFLRRAGKQDQTSNNSDDKNRNSSEERSRPN 123
 DB 67 HAHKEEMIKINKGRFOQAVDAVEEFLRRAGKQDQTSNNSDDKNRNSSEERSRPN 126
 QY 124 DEKKLSLSEKGVKKNMGEKKRVSSESSSRRGSKSPKRAQEQSPKRGKPPKDEKDLTI 183
 DB 127 DEKKLSLSEKGVKKNMGEKKRVSSESSSRRGSKSPKRAQEQSPKRGKPPKDEKDLTI 186
 QY 184 PESSTVKGMMAGPMAAFKWOPTASEPVKADPHFHFLLSQTEKPAVCYQAITKKLICE 243
 DB 187 PESSTVKGMMAGPMAAFKWOPTASEPVKADPHFHFLLSQTEKPAVCYQAITKKLICE 246
 QY 244 EETGSTSIOAADSTAVNGSITPTDKKIGFLGLGMSGIVSNLLKMGHTVTVNRTAEK 303
 DB 247 EETGSTSIOAADSTAVNGSITPTDKKIGFLGLGMSGIVSNLLKMGHTVTVNRTAEK 305
 QY 304 DLFIQEGARLGRTPAEVWSTCDITFACVSDPKAAKDLVLGSPGVLOIRPGKCYVDMSTV 363
 DB 306 -----EGARLGRTPAEVWSTCDITFACVSDPKAAKDLVLGSPGVLOIRPGKCYVDMSTV 360
 QY 364 DADTVTELAQVIVSRGRFLEAPVSGNQOLSDNGLMVLVLAAGDRGLYEDSCSCFOAMGKT 423
 DB 361 DADTVTELAQVIVSRGRFLEAPVSGNQOLSDNGLMVLVLAAGDRGLYEDSCSCFOAMGKT 420
 QY 424 SPFLGEVGNAAKMLIVNVVQGSFMATIAEGLTFLAHVTGSGQOTLLDILNQGSLAFILD 483
 DB 421 SPFLGEVGNAAKMLIVNVVQGSFMATIAEGLTFLAHVTGSGQOTLLDILNQGSLAFILD 480
 QY 484 QXCONILQGNFKPDPFYLKVIQKDLRLAIALGDVNVHPTMAAANVEYKRAKALDQSDND 543
 DB 481 QXCONILQGNFKPDPFYLKVIQKDLRLAIALGDVNVHPTMAAANVEYKRAKALDQSDND 540
 QY 544 MSAYVRYATH 553
 DB 541 MSAYVRYATH 550
 RESULT 5
 ADC46331
 ID ADC46331 standard; protein; 550 AA.
 XX AC ADC46331;
 XX DT 18-DEC-2003 (first entry)
 XX DE Human neoplastic disease-associated gene 123 protein #1.
 XX KW Neoplastic disease-associated polypeptide; gene therapy;
 KW hyperproliferative disease; cancer; autoimmune disorder; diabetes;
 KW rheumatoid arthritis; systemic lupus erythematosus; multiple sclerosis;
 KW autoimmune thyroiditis; haemolytic anaemia; haematopoietic disorder;
 KW haematologic disorder; anaemia; thrombocytopenia; allergic reaction;
 KW asthma; ezema; inflammatory disorder; ischaemia-reperfusion injury;
 KW inflammatory bowel disease; Crohn's disease; neurodegenerative disorder;
 KW Alzheimer's disease; Parkinson's disease; renal disorder;
 KW acute glomerulonephritis; end-stage renal disease;
 KW cardiovascular disorder; atherosclerosis; myocarditis;
 KW infectious disease; AIDS; cachexia; anorexia; wound healing;
 KW epithelial cell proliferation; Human.
 XX OS Homo sapiens.
 XX PN US2003082758-A1.
 XX XX

Thu Sep 16 09:36:51 2004

PD 01-MAY-2003.
XX 22-MAR-2002; 2002US-00103313.
PF 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239355P.
PR 13-OCT-2000; 2000US-0239377P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2000US-0259678P.
PR 17-JAN-2001; 2001US-00764854.
XX
(HUMA-) HUMAN GENOME SCI INC.
XX
Rosen CA, Ruben SM, Barash SC;
XX
XX

DR WPI: 2003-786918/74.
 DR N-PSDB; ADC46047.
 XX
 PT New isolated human neoplastic disease-associated polypeptides and
 PT polynucleotides, useful for diagnosing, preventing, prognosticating or
 PT treating medical conditions such as cancer, AIDS, diabetes or Parkinson's
 PT disease.
 XX
 PS Claim 11; SEQ ID NO 417; 302pp; English.
 XX

CC The invention relates to one of 238 disclosed human neoplastic disease-
 CC associated polypeptides encoded by 171 disclosed cDNA sequences
 CC (including their domains, epitopes, full-length proteins, allelic variants
 CC or species homologues). Also included are there encoding nucleic acids, a
 CC recombinant vector comprising the nucleic acid, a recombinant host cell
 CC comprising the nucleic acid (expressing the protein), an isolated
 CC antibody that binds specifically to the isolated polypeptide, preventing,
 CC treating or ameliorating a medical condition, diagnosing a pathological
 CC condition or a susceptibility to a pathological condition, a pathological
 CC identifying a binding partner to the polypeptide, identifying an activity
 CC in a biological assay, and the gene corresponding to the cDNA sequence.
 CC The polypeptides, polynucleotides and antibodies are useful for
 CC detecting, preventing, diagnosing, prognosticating, treating or
 CC ameliorating medical conditions such as hyperproliferative diseases or
 CC cancer, autoimmune disorders (e.g. diabetes, rheumatoid arthritis,
 CC systemic lupus erythematosus, multiple sclerosis, autoimmune thyroiditis
 CC or haemolytic anaemia), haematopoietic or haematologic disorders (e.g.
 CC anaemia or thrombocytopaenia), allergic reactions including asthma or
 CC eczema, inflammatory disorders (e.g. ischaemia-reperfusion injury,

Query Match 97.6%; Score 2798; DB 7; Length 550;
 Best Local Similarity 98.7%; Pred. No. 9.3e-249;
 Matches 543; Conservative 0; Mismatches 1; Indels 6; Gaps 1;
 QY 4 VSLRLGDLVWGLGRYPWPVKIYNPPKDLKPRGKKCFVFFGTEHAWIKVEQLKPY 63
 Db 7 VSLRLGDLVWGLGRYPWPVKIYNPPKDLKPRGKKCFVFFGTEHAWIKVEQLKPY 66
 QY 64 HAHKEMIKINKGRFQQAQVDAVEEFLLRAKQDQTSNHSDDKNNRNSSEERSRPNSG 123
 Db 67 HAHKEMIKINKGRFQQAQVDAVEEFLLRAKQDQTSNHSDDKNNRNSSEERSRPNSG 126
 QY 124 DEKRKLISLSEGVKKMGKGRVSSGSSGERSKSPKRAQSGSPKRPKDEKDLTI 183
 Db 127 DEKRKLISLSEGVKKMGKGRVSSGSSGERSKSPKRAQSGSPKRPKDEKDLTI 186
 QY 184 PSSSTVKGMMAGPMAAFKWOPTASEPVKDADPHFHFLLSOTKPAVCYQAITKKLKICE 243
 Db 187 PSSSTVKGMMAGPMAAFKWOPTASEPVKDADPHFHFLLSOTKPAVCYQAITKKLKICE 246
 QY 244 BETGSTSIQAADSTAVNGSITPTDKKIGFLGLMGSGIVSNLLKMGHTVTVMNRTAEC 303
 Db 247 BETGSTSIQAADSTAVNGSITPTDKKIGFLGLMGSGIVSNLLKMGHTVTVMNRTAEC 305
 QY 304 DLFIQEGARLGRTAEVNSTCDITFACVSDPKAAKDLVLGPGSVLQGIKPGCYVDMSTV 363
 Db 306 -----EGARLGRTAEVNSTCDITFACVSDPKAAKDLVLGPGSVLQGIKPGCYVDMSTV 360
 QY 364 DADVTTELAQVIVSRGGRFLEAPVSGNQLSNDGMLVILAGDGLGYEDCSCFOAMGKT 423
 Db 361 DADVTTELAQVIVSRGGRFLEAPVSGNQLSNDGMLVILAGDGLGYEDCSCFOAMGKT 420
 QY 424 SFFLGEVGNAAKMLIVNVQGSFMATIAEGITLAHVTGQSQTLLDILNOGLASIFLD 483
 Db 421 SFFLGEVGNAAKMLIVNVQGSFMATIAEGITLAHVTGQSQTLLDILNOGLASIFLD 480
 QY 484 QKCNILQGNFKPDYLYKIQKDLALALGDVAVNHPTPMAAANEVYKRAKALQDSND 543
 Db 481 QKCNILQGNFKPDYLYKIQKDLALALGDVAVNHPTPMAAANEVYKRAKALQDSND 540
 QY 544 MSAYVRYIYH 553
 Db 541 MSAYVRYIYH 550

RESULT 6
 AAE38240
 ID AAE38240 standard; protein; 472 AA.
 XX
 AC AAE38240;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Human enzyme (ENZM) protein #32.
 XX
 KW Human; enzyme; ENZM; immune disorder; infection; myocardial infarction;
 KW gene therapy; anaemia; acquired immune deficiency syndrome; infection;
 KW reproductive disorder; cardiovascular; eye; cell proliferation; cancer;
 KW AIDS; allergy; asthma; Addison's disease; diabetes; goitre; impotence;
 KW infertility; atherosclerosis; metabolic disorder.
 XX
 OS Homo sapiens.
 XX
 PN W02003052075-A2.
 XX
 PD 26-JUN-2003.
 XX
 PF 12-DEC-2002; 2002WO-US040161.
 XX
 PR 14-DEC-2001; 2001US-0340357P.
 PR 20-DEC-2001; 2001US-0342962P.
 PR 21-DEC-2001; 2001US-034358P.
 PR 22-JAN-2002; 2002US-0351107P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Chawla NK, Lee SY, Ring HZ, Lee EA, Forsythe IJ, Khare R;
 PI Tran UK, Kable AE, Richardson TW, Emerling BM, Lindquist EA;
 PI Baughn MR, Hafalia AJA, Jin P, Swarnakar A, Li JX, Marquis JP;
 PI Lee S, Gorvad AE, Sprague WW, Becha SD, Elliott VS;
 XX
 DR WPI: 2003-533016/50.
 DR N-PSDB; AAD57520.
 XX
 PT New human enzyme (ENZM) polypeptide, useful for preparing a composition
 PT for treating a disease associated with decreased expression or
 PT overexpression of ENZM e.g. cancer.
 XX
 PS Claim 1; Page 298-299; 264pp; English.
 XX
 CC The invention relates to human enzyme (ENZM) polypeptides and their
 CC corresponding polynucleotides. ENZM sequences are useful for preparing a
 CC composition for diagnosing or treating a disease or condition associated
 CC with decreased expression or overexpression of functional ENZM. The
 CC disorders include immune disorders (anaemia, allergy or asthma),
 CC infectious disorders (viral, fungal, parasitic or protozoal infection),
 CC immune deficiencies (acquired immune deficiency syndrome; AIDS),
 CC metabolic disorder (Addison's disease, diabetes or goitre), reproductive
 CC disorders (infertility or impotence), cardiovascular disorders
 CC (atherosclerosis or myocardial infarction), eye disorders and cell
 CC proliferative disorders (cancer). ENZM DNA is useful in gene therapy. The
 CC present sequence is human ENZM protein
 XX
 SQ Sequence 472 AA;

Query Match 83.6%; Score 2396.5; DB 6; Length 472;
 Best Local Similarity 85.2%; Pred. No. 8.1e-212;
 Matches 471; Conservative 0; Mismatches 1; Indels 81; Gaps 1;
 QY 1 MAAYSLRLGDLVWGLGRYPWPVKIYNPPKDLKPRGKKCFVFFGTEHAWIKVEQL 60
 Db 1 MAAYSLRLGDLVWGLGRYPWPVKIYNPPKDLKPRGKKCFVFFGTEHAWIKVEQL 60
 QY 61 KPYHAHKEEMIKINKGRFQQAQVDAVEEFLLRAKQDQTSNHSDDKNNRNSSEERSRP 120
 Db 61 KPYHAHKEEMIKINKGRFQQAQVDAVEEFLLRAKQDQTSNHSDDKNNRNSSEERSRP 120
 ----- 98

Thu Sep 16 09:36:51 2004

us-10-067-482-2.rag

QY	121	NSGDEKRKLSLSEGVKKQKMGEGKRVKSSGSSERGSKSLPKRAEQOSPKRGPRPKDQXD	180
Db	99	-----D	99
QY	181	LITPESSTVKGMMAGMAAFKQPTASEPVKQADPHFHFLLSOTKPAVCYQAIKKLK	240
Db	100	LITPESSTVKGMMAGMAAFKQPTASEPVKQADPHFHFLLSOTKPAVCYQAIKKLK	159
QY	241	ICBEETGTSIOAADSTAVNGSITPTDKKIGFLGLGMSGGIVSNLLKMGHTVTVNRTA	300
Db	160	ICBEETGTSIOAADSTAVNGSITPTDKKIGFLGLGMSGGIVSNLLKMGHTVTVNRTA	219
QY	301	EKCDLFIQEGARLGRTPAEVWSTCDITFACVSPKAAKDLVLPSPGVQIRPKCYVDM	360
Db	220	EKCDLFIQEGARLGRTPAEVWSTCDITFACVSPKAAKDLVLPSPGVQIRPKCYVDM	279
QY	361	STVDADTVTELAQVTVSRGRFLEAPVSGNQQLSNDGMLVILAAGDRGLYEDCSSCFQAM	420
Db	280	STVDADTVTELAQVTVSRGRFLEAPVSGNQQLSNDGMLVILAAGDRGLYEDCSSCFQAM	339
QY	421	GKTSFFLGEVGNAAKMMLIYNNVQGSFMTIAEGLTLAHTVQSGQOQLLILNQGLASI	480
Db	340	GKTSFFLGEVGNAAKMMLIYNNVQGSFMTIAEGLTLAHTVQSGQOQLLILNQGLASI	399
QY	481	FLDQKQCNILQGNPKDFYLYKIQKDLRLAIALGDVAVNHTPMAAAANEVYKRAKALDQS	540
Db	400	FLDQKQCNILQGNPKDFYLYKIQKDLRLAIALGDVAVNHTPMAAAANEVYKRAKALDQS	459
QY	541	DNDMSAVYRAYIH 553	
Db	460	DNDMSAVYRAYIH 472	
RESULT 7			
AAU20515			
ID	AAU20515 standard; protein; 269 AA.		
XX	AAU20515;		
AC	AAU20515;		
XX	AAU20515;		
DT	06-DEC-2001 (first entry)		
XX	Human secreted protein, Seq ID No 507.		
DE	Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;		
XX	rheumatoid arthritis; antiarteriosclerotic; cardiac; vascular;		
KW	cerebroprotective; thrombolytic; antimicrobial; ophthalmological;		
KW	cytostatic; Alzheimer's disease; Parkinson's disease; human; cancer;		
KW	multiple sclerosis; cancer; hyperproliferative disorder; infection;		
KW	Gaucher's disease; neurological disease; cerebrovascular disorder;		
XX	thrombosis; wound healing.		
OS	Homo sapiens.		
XX	WO200155326-A2.		
PN	02-AUG-2001.		
PD	17-JAN-2001; 2001WO-US001347.		
XX	31-JAN-2000; 2000US-0179065P.		
XX	04-FEB-2000; 2000US-0180628P.		
PR	24-FEB-2000; 2000US-0184664P.		
PR	02-MAR-2000; 2000US-0186350P.		
PR	16-MAR-2000; 2000US-0189874P.		
PR	17-MAR-2000; 2000US-0190076P.		
PR	18-APR-2000; 2000US-0198123P.		
PR	19-MAY-2000; 2000US-0205515P.		
PR	07-JUN-2000; 2000US-0209467P.		
PR	28-JUN-2000; 2000US-0214886P.		
PR	30-JUN-2000; 2000US-0215135P.		
PR	07-JUL-2000; 2000US-0216647P.		
PR	07-JUL-2000; 2000US-0216880P.		
PR	11-JUL-2000; 2000US-0217487P.		
PR	11-JUL-2000; 2000US-0217496P.		
PR	14-JUL-2000; 2000US-0218290P.		
PR	26-JUL-2000; 2000US-0220963P.		
PR	26-JUL-2000; 2000US-0220964P.		
PR	14-AUG-2000; 2000US-0224518P.		
PR	14-AUG-2000; 2000US-0224519P.		
PR	14-AUG-2000; 2000US-0225213P.		
PR	14-AUG-2000; 2000US-0225214P.		
PR	14-AUG-2000; 2000US-0225266P.		
PR	14-AUG-2000; 2000US-0225267P.		
PR	14-AUG-2000; 2000US-0225268P.		
PR	14-AUG-2000; 2000US-0225270P.		
PR	14-AUG-2000; 2000US-0225447P.		
PR	14-AUG-2000; 2000US-0225757P.		
PR	14-AUG-2000; 2000US-0225758P.		
PR	14-AUG-2000; 2000US-0225759P.		
PR	14-AUG-2000; 2000US-0226279P.		
PR	22-AUG-2000; 2000US-0226681P.		
PR	22-AUG-2000; 2000US-0226688P.		
PR	22-AUG-2000; 2000US-0227182P.		
PR	23-AUG-2000; 2000US-0227009P.		
PR	30-AUG-2000; 2000US-0228924P.		
PR	01-SEP-2000; 2000US-0229287P.		
PR	01-SEP-2000; 2000US-0229343P.		
PR	01-SEP-2000; 2000US-0229344P.		
PR	01-SEP-2000; 2000US-0229345P.		
PR	05-SEP-2000; 2000US-0229509P.		
PR	05-SEP-2000; 2000US-0229513P.		
PR	06-SEP-2000; 2000US-0230437P.		
PR	06-SEP-2000; 2000US-0230438P.		
PR	08-SEP-2000; 2000US-0231242P.		
PR	08-SEP-2000; 2000US-0231243P.		
PR	08-SEP-2000; 2000US-0231244P.		
PR	08-SEP-2000; 2000US-0231413P.		
PR	08-SEP-2000; 2000US-0231414P.		
PR	08-SEP-2000; 2000US-0232080P.		
PR	08-SEP-2000; 2000US-0232081P.		
PR	12-SEP-2000; 2000US-0231968P.		
PR	14-SEP-2000; 2000US-0232397P.		
PR	14-SEP-2000; 2000US-0232398P.		
PR	14-SEP-2000; 2000US-0232399P.		
PR	14-SEP-2000; 2000US-0232400P.		
PR	14-SEP-2000; 2000US-0232401P.		
PR	14-SEP-2000; 2000US-0233063P.		
PR	14-SEP-2000; 2000US-0233064P.		
PR	14-SEP-2000; 2000US-0233065P.		
PR	21-SEP-2000; 2000US-0234223P.		
PR	21-SEP-2000; 2000US-0234274P.		
PR	25-SEP-2000; 2000US-0234997P.		
PR	25-SEP-2000; 2000US-0234998P.		
PR	26-SEP-2000; 2000US-0235484P.		
PR	27-SEP-2000; 2000US-0235834P.		
PR	27-SEP-2000; 2000US-0235836P.		
PR	29-SEP-2000; 2000US-0236327P.		
PR	29-SEP-2000; 2000US-0236367P.		
PR	29-SEP-2000; 2000US-0236368P.		
PR	29-SEP-2000; 2000US-0236369P.		
PR	29-SEP-2000; 2000US-0236370P.		
PR	02-OCT-2000; 2000US-0236802P.		
PR	02-OCT-2000; 2000US-0237037P.		
PR	02-OCT-2000; 2000US-0237038P.		
PR	02-OCT-2000; 2000US-0237039P.		
PR	02-OCT-2000; 2000US-0237040P.		
PR	13-OCT-2000; 2000US-0239935P.		
PR	13-OCT-2000; 2000US-0239937P.		
PR	20-OCT-2000; 2000US-0240960P.		
PR	20-OCT-2000; 2000US-0241221P.		
PR	20-OCT-2000; 2000US-0241785P.		
PR	20-OCT-2000; 2000US-0241786P.		
PR	20-OCT-2000; 2000US-0241787P.		
PR	20-OCT-2000; 2000US-0241808P.		
PR	20-OCT-2000; 2000US-0241809P.		

20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246479P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 08-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 (HUMA-) HUMAN GENOME SCI INC.
 PA Rosen CA, Barash SC, Ruben SM;
 PI WPI: 2001-451931/48.
 XX N-PSDB; AAS33224.
 DR New nucleic acids and polypeptides, useful for diagnosing, preventing or
 PT treating medical conditions.
 XX Claim 11; SEQ ID NO 507; 753pp; English.
 PS The invention relates to novel isolated nucleic acid molecules (I)
 CC encoding human secreted proteins (II). (I) and (II) are used to prevent,
 CC treat or ameliorate a medical condition in e.g. humans, mice, rabbits,
 CC goats, horses, cats, dogs, chickens or sheep. (I) and (II) may be used in
 CC the prevention, treatment and diagnosis of diseases associated with
 CC inappropriate expression of secreted proteins. (I) and complementary
 CC sequences may also be used as DNA probes in diagnostic assays (e.g.
 CC polymerase chain reactions (PCR)) to detect and quantitate the presence
 CC of similar nucleic acid sequences in samples, and so which patients may
 CC be in need of restorative therapy. (II) may also be used as antigens in

the production of antibodies and in assays to identify modulators
 (agonists and antagonists) of the expression and activity of the secreted
 proteins. The anti-(ii) antibodies and antagonists may also be used to
 down regulate expression and activity of (ii). The anti-(ii) antibodies
 may also be used as diagnostic agents for detecting the presence of (ii)
 in samples (e.g. by enzyme linked immunosorbant assay (ELISA)). The
 disorders include for example: immune/ autoimmune diseases (e.g. HIV
 (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis
 and multiple sclerosis), cancers and hyperproliferative disorders (e.g.
 melanomas, neoplasms of the breast or liver, Sezary syndrome and
 Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,
 Parkinson's disease and Charcot-Marie-Tooth disease), cardio-/
 cerebrovascular disorders (e.g. cardiac arrest, tachycardia, angina and
 thrombosis), infections caused by bacteria, viruses and fungi and ocular
 disorders (e.g. corneal infections). (i) and (ii), agonists, antagonists
 and antibodies can also be used to promote wound healing, maintain organs
 CC before transplantation, and support cell culture of primary tissues.

Query Match 49.0%; Score 1404; DB 4; Length 269;
 Best Local Similarity 99.3%; Pred. No. 1.3e-120;
 Matches 267; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 4 VSLRLGDLVWGLGRYPWPFGKIVNPPKDKLKKPRGKCFVKKFGTGDHAWIKVEQLKPY 63
 Db 1 VSLXGLDLVWGLGRYPWPFGKIVNPPKDKLKKPRGKCFVKKFGTGDHAWIKVEQLKPY 60
 QY 64 HAHKEWIKINKGRFQQAQVDAVEEFLLRAKQDQTSNHSDDKRRNSSESRSPNSG 123
 Db 61 HAHKEWIKINKGRFQQAQVDAVEEFLLRAKQDQTSNHSDDKRRNSSESRSPNSG 120
 QY 124 DEKRKLSLSEGVKKWNGEGKKEVSSGSSERGSKSPKRAQEQSPKRGPPDKDLTI 183
 Db 121 DEKRKLSLSEGVKKWNGEGKKEVSSGSSERGSKSPKRAQEQSPKRGPPDKDLTI 180
 QY 184 PESSTVKGMMWAGPMAAFKMQPTASEPVDADPHFHFLLSQTEKPAVCYQAITKKLKICE 243
 Db 181 PESSTVKGMMWAGPMAAFKMQPTASEPVDADPHFHFLLSQTEKPAVCYQAITKKLKICE 240
 QY 244 EETGSTSIOAADSTAVNGSITPTDKKIGF 272
 Db 241 EETGSTSIOAADSTAVNGSITPTDKKIGF 269

RESULT 8

AAU21813
 ID AAU21813 standard; protein; 269 AA.

AC AAU21813;

XX 06-DEC-2001 (first entry)

XX Novel human neoplastic disease associated polypeptide #246.

XX Human; neoplastic disease associated polypeptide; cancer;
 KW hyperproliferative disorder; neural disorder; immune system disorder;
 KW muscular disorder; reproductive disorder; gastrointestinal disorder;
 KW pulmonary disorder; cardiovascular disorder; renal disorder;
 KW neuroprotective; cytostatic; anti inflammatory; vasotropic.

OS Homo sapiens.

XX WO200155163-A1.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US001358.

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180628P.

XX 24-FEB-2000; 2000US-0184664P.

XX 02-MAR-2000; 2000US-0186350P.

XX 16-MAR-2000; 2000US-0189874P.

XX 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225366P.
PR 14-AUG-2000; 2000US-0225367P.
PR 14-AUG-2000; 2000US-0225368P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 05-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0239353P.
PR 13-OCT-2000; 2000US-0239377P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241926P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 05-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Rosen CA, Barash SC, Ruben SM;
PI
XX WPI; 2001-465558/50.
XX N-PSDB; AAS35012.
XX Novel polypeptides and polynucleotides useful as diagnostic reagents to
PT diagnose diseases or disorders associated with aberrant expression or
PT activity of polypeptides, and for treating cancers, rheumatoid arthritis.
XX
PS Claim 11; SEQ ID NO 540; 687pp; English.
XX The present invention relates to the isolation of novel human neoplastic
CC disease associated polypeptides, and cDNA (AAS34767-AAS35050) and DNA

CC sequences encoding for these polypeptides. The sequences of the invention
 CC are useful in the diagnosis, treatment, prevention and/or prognosis of
 CC disorders involving neoplastic disease such as hyperproliferative
 CC disorders (e.g. leukemia, bone cancer, bladder cancer, brain stem
 CC glioma, adult liver cancer, childhood cerebellar astrocytoma, or
 CC Hodgkin's lymphoma). The sequences of the invention may also be useful
 CC for treating other disorders such as neural disorders, immune system
 CC disorders, muscular disorders, reproductive disorders, gastrointestinal
 CC disorders, pulmonary disorders, cardiovascular disorders and renal
 CC disorders. The polynucleotide sequences of the invention are also useful
 CC in gene therapy. AAU21568-AAU21851 represent the novel human neoplastic
 CC disease associated polypeptides of the invention. Note: The sequence data
 CC for this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX

SQ Sequence 269 AA;

Query Match 49.0%; Score 1404; DB 4; Length 269;
 Best Local Similarity 99.3%; Pred. No. 1.3e-120;
 Matches 267; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 4 VSLRLGDLVWGLGRYPWPFGKIVNPPKDLKPRGKKCFVKECTEDHAWIKVEQLKPY 63
 DB 1 VSLRLGDLVWGLGRYPWPFGKIVNPPKDLKPRGKKCFVKECTEDHAWIKVEQLKPY 60
 QY 64 HAHKEMIKINKGRFQQAQVDAVEEFLRAKGDQTSNNSSDDKNRNSSEERSRPNNG 123
 DB 61 HAHKEMIKINKGRFQQAQVDAVEEFLRAKGDQTSNNSSDDKNRNSSEERSRPNNG 120
 QY 124 DEKRKLISLSEGVKVMGKGRVSVSGSSERGSKSPKRAEQSPKRGPRPPKDEKDLTI 183
 DB 121 DEKRKLISLSEGVKVMGKGRVSVSGSSERGSKSPKRAEQSPKRGPRPPKDEKDLTI 180
 QY 184 PSSVTVKGMAGPMAAFKWPQTASEPVKDADPHFHHLLSOTEKPAVCYQAITKKLKICE 243
 DB 181 PSSVTVKGMAGPMAAFKWPQTASEPVKDADPHFHHLLSOTEKPAVCYQAITKKLKICE 240
 QY 244 EPTGTSIQAADSTAVNGSIITPTDKKIGF 272
 DB 241 EPTGTSIQAADSTAVNGSIITPTDKKIGF 269

RESULT 9

ID ADC46454
 AC ADC46454 standard; protein; 269 AA.

XX ADC46454;

XX 18-DEC-2003 (first entry)

DE Human neoplastic disease-associated gene 123 protein #2.

XX Neoplastic disease-associated polypeptide; gene therapy;
 KW hyperproliferative disease; cancer; autoimmune disorder; diabetes;
 KW rheumatoid arthritis; systemic lupus erythematosus; multiple sclerosis;
 KW autoimmune thyroiditis; haemolytic anaemia; haematopoietic disorder;
 KW haematologic disorder; anaemia; thrombocytopaenia; allergic reaction;
 KW asthma; eczema; inflammatory disorder; ischaemia-reperfusion injury;
 KW inflammatory bowel disease; Crohn's disease; renal disorder;
 KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
 KW acute glomerulonephritis; end-stage renal disease;
 KW cardiovascular disorder; atherosclerosis; myocarditis;
 KW infectious disease; AIDS; cachexia; anorexia; wound healing;
 KW epithelial cell proliferation; Human.

OS Homo sapiens.

XX US2003082758-A1.

XX 01-MAY-2003.

XX 22-MAR-2002; 2002US-00103313.

XX 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226688P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0228287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 21-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 25-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 26-SEP-2000; 2000US-0234998P.
 PR 27-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.

PR 02-OCT-2000; 2000US-0235802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-0076485A.
(HUMA-) HUMAN GENOME SCI INC.
XX PA
XX PI Rosen CA, Ruben SM, Barash SC;
XX DR WPI; 2003-786918/74.
XX N-PSDB; ADC46170.
XX

PT New isolated human neoplastic disease-associated polypeptides and
PT polynucleotides, useful for diagnosing, preventing, prognosticating or
PT treating medical conditions such as cancer, AIDS, diabetes or Parkinson's
disease.
XX Claim 11; SEQ ID NO 540; 302pp; English.
XX The invention relates to one of 238 disclosed human neoplastic disease-
associated polypeptides encoded by 171 disclosed cDNA sequences
(including their domains, epitopes, full-length proteins, allelic variants
or species homologues). Also included are there encoding nucleic acids, a
recombinant vector comprising the nucleic acid, a recombinant host cell
comprising the nucleic acid (expressing the protein), an isolated
antibody that binds specifically to the isolated polypeptide, preventing,
treating or ameliorating a medical condition, diagnosing a pathological
condition or a susceptibility to a pathological condition in a subject,
identifying a binding partner to the polypeptide, identifying an activity
in a biological assay, and the gene corresponding to the cDNA sequence.
CC The polypeptides, polynucleotides and antibodies are useful for
CC detecting, preventing, diagnosing, prognosticating, treating or
CC ameliorating medical conditions such as hyperproliferative diseases or
cancer, autoimmune disorders (e.g. diabetes, rheumatoid arthritis,
systemic lupus erythematosus, multiple sclerosis, autoimmune thyroiditis
or haemolytic anaemia), haematopoietic or haematologic disorders (e.g.
anaemia or thrombocytopaenia), allergic reactions including asthma or
eczema, inflammatory disorders (e.g. ischaemia-reperfusion injury,
Query Match 49.0%; Score 1404; DB 7; Length 269;
Best Local Similarity 99.3%; Pred. No. 1.3e-120;
Matches 267; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4 VSLRLGDLVWGKLGKRYPPFPGKIVNPKOLKPRGKCKPFVFGTGHAWIKVEQLKPY 63
DB 1 VSLXLDLVWGKLGKRYPPFPGKIVNPKDLKPRGKCKPFVFGTGHAWIKVEQLKPY 60
QY 64 HAHKEEMIKNGKRFQOAVDAVEEFLRAKGDQTSNHNSSDDKNRNSSEERSRNSG 123
DB 61 HAHKEEMIKNGKRFQOAVDAVEEFLRAKGDQTSNHNSSDDKNRNSSEERSRNSG 120
QY 124 DEKXKLSLSEKGVKKNMGEGKRVSSGSSRGSKSPKRAQEQSPFRKGRPPPKDKDLTI 183
DB 121 DEKRLSLSEKGVKKNMGEGKRVSSGSSRGSKSPKRAQEQSPFRKGRPPPKDKDLTI 180
QY 184 PESSTVKGMMAGPMAAFKQPTASEPVKADADPHPHFLLSOTEPKPAVCYQAITKKLKICE 243
DB 181 PESSTVKGMMAGPMAAFKQPTASEPVKADADPHPHFLLSOTEPKPAVCYQAITKKLKICE 240
QY 244 EETGSTSIQAADSTAVNGSITPTDKKIGF 272
DB 241 EETGSTSIQAADSTAVNGSITPTDKKIGF 269
RESULT 10
AAW09404
ID AAW09404 standard; protein; 249 AA.
XX AC AAW09404;
XX XX
DT 20-MAY-1997 (first entry)
XX DE Human hepatoma-derived growth factor (HDGF-2).
XX KW Hepatoma-derived growth factor-2; HDGF-2; chemokine; wound healing;
KW vulnery; burn; ulcer; thrombosis; arteriosclerosis; antibody;
KW antagonist; tumour; therapy.
XX OS Homo sapiens.
XX XX
XX PN WO9639485-A1.
XX XX
XX PD 12-DEC-1996.
XX XX
XX PF 05-JUN-1995; 95WO-US006731.

XX 05-JUN-1995; 95WO-US0006731.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Kunsch CA, Rosen CA;
 PI WPI; 1997-043108/04.
 DR N-PSDB; AAT47520.
 XX
 XX DNA encoding human-derived growth factor polypeptide - useful to promote
 PT wound healing as result of burns or ulcers.
 XX
 XX Claim 7; Fig 2; 54pp; English.
 XX
 CC Human hepatoma-derived growth factor-2 (HDGF-2) (AAW09404) shows 23%
 CC identity and 61% similarity over a 201-amino acid stretch to human HDGF,
 CC a mitogen for fibroblasts. Its amino acid sequence was deduced from a
 CC cDNA clone (AAT47520) isolated from a human umbilical vein endothelial
 CC tissue cDNA library. Recombinant HDGF-2 can be produced in transformed
 CC host (e.g. E. coli, Sf9, COS) cells. It is useful in stimulating tissue
 CC repair and tissue growth and can also be used to produce antibodies and
 CC to screen for antagonists that retard tumour growth and reduce scarring
 XX
 SQ Sequence 249 AA;
 Query Match 45.0%; Score 1291; DB 2; Length 249;
 Best Local Similarity 99.6%; Pred. No. 3e-110;
 Matches 243; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAASVLRGLDVLVWGLGRYPWPFGKIVNPPKOLKPRGKKCFVVKFGTDEHAWIKVEQL 60
 Db 1 MAASVLRGLDVLVWGLGRYPWPFGKIVNPPKOLKPRGKKCFVVKFGTDEHAWIKVEQL 60
 QY 61 KPYHAHKEEMIKINKGRFQQAQVDAVEEFLRRAKGDQTSNHSDDKNNRNSSEERSRP 120
 Db 61 KPYHAHKEEMIKINKGRFQQAQVDAVEEFLRRAKGDQTSNHSDDKNNRNSSEERSRP 120
 QY 121 NSGDEKRLSLSEGKVKGMGKGRVSSGSSGSRGSKSLKRAQESPKRGRPPPKDEKD 180
 Db 121 NSGDEKRLSLSEGKVKGMGKGRVSSGSSGSRGSKSLKRAQESPKRGRPPPKDEKD 180
 QY 181 LTIPESSIVKGMAGPMFAFKWQPTASEPVKQADPHFHFLLSQTEKPAVCYQAITKKLK 240
 Db 181 LTIPESSIVKGMAGPMFAFKWQPTASEPVKQADPHFHFLLSQTEKPAVCYQAITKKLK 240
 QY 241 ICEE 244
 Db 241 ICED 244
 RESULT 11
 ABU09982
 ID ABU09982 standard; protein; 249 AA.
 XX
 AC ABU09982;
 XX
 DT 30-JUL-2003 (first entry)
 XX
 DE Human hepatoma-derived growth factor 2 (HDGF-2).
 XX
 KW Human; hepatoma-derived growth factor-2; HDGF-2; vasotropic; vulnery;
 KW neurotropic; neuroprotective; antiparkinsonian; cytostatic;
 KW antiinflammatory; gene therapy; vascular endothelial cell growth;
 KW re-vascularisation; ischaemic tissue; thrombosis; arteriosclerosis;
 KW cardiovascular condition; mesodermal induction; limb regeneration;
 KW wound healing; injury; burn; surgery; ulcer; neuronal growth;
 KW neuronal damage; neuronal disorder; neuro-degenerative condition;
 KW Alzheimer's disease; Parkinson's disease; AIDS-related complex;
 KW chondrocyte growth; bone regeneration; periodontal regeneration;
 KW tissue transplant; bone graft; skin aging; sunburn; keratinocyte growth;
 KW hair loss; hair-forming cell activation; melanocyte growth;
 KW haematopoietic cell growth; haematopoietic cell differentiation;

KW bone marrow cell; organ transplantation; mesodermal tissue; cell growth;
 KW cell proliferation; tumour; hypervascular disease;
 KW epithelial lens cell proliferation; extracapsular cataract surgery;
 KW retinosis; balloon angioplasty; inflammation; scar tissue.
 XX
 XX Homo sapiens.
 XX OS
 XX US2003022312-A1.
 XX
 XX 30-JAN-2003.
 XX
 XX 15-NOV-2001; 2001US-00987755.
 XX
 XX 05-JUN-1995; 95US-00464600.
 XX 05-MAR-1999; 99US-00263625.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Kunsch CA, Rosen CA;
 PI WPI; 2003-447425/42.
 DR N-PSDB; ACA61586.
 XX

Novel isolated human hepatoma-derived growth factor-2 polypeptide useful

for promoting healing in wounds, treating or prevention of Alzheimer's
 disease, stimulating chondrocyte growth, preventing skin aging, and
 preventing hair loss.

Claim 2; Fig 1; 21pp; English.

The invention describes a polypeptide (I) having a deduced amino acid
 sequence of human hepatoma-derived growth factor-2 (HDGF-2) polypeptide
 sequence of 249 amino acids (S2) as given in the specification, or its
 fragments, analogues or derivatives. (I) is useful for stimulating
 vascular endothelial cell growth, and thus is employed in treating for
 stimulating re-vascularisation of ischaemic tissues due to various
 disease conditions such as thrombosis, arteriosclerosis, and other
 cardiovascular conditions; stimulating mesodermal induction and limb
 regeneration in early embryos; promoting healing in wounds due to
 injuries, burns, surgery, and ulcers; stimulating neuronal growth, and
 thus for treating or preventing neuronal damage occurring in certain
 neuronal disorders or neuro-degenerative conditions such as Alzheimer's
 disease, Parkinson's disease, and AIDS-related complex; stimulating
 chondrocyte growth and thus for enhancing bone and periodontal
 regeneration, and aiding in tissue transplants or bone grafts; preventing
 skin aging due to sunburn by stimulating keratinocyte growth; preventing
 hair loss by activating hair-forming cells and promoting melanocyte
 growth; stimulating growth and differentiation of haematopoietic cells
 and bone marrow cells; maintaining organs before transplantation; and
 inducing tissue of mesodermal origin to differentiate in early embryos.
 The antagonists of (I) are useful for inhibiting the cell growth and
 proliferation effects of the polypeptides on neoplastic cells and
 tissues, and therefore are useful for preventing abnormal cellular growth
 and proliferation e.g. in tumour formation or growth; preventing
 hypervascular diseases and preventing the proliferation of epithelial
 lens cells after extracapsular cataract surgery; preventing retinosis
 after balloon angioplasty; and preventing inflammation and growth of scar
 tissue during wound healing. This is the amino acid sequence of human
 hepatoma-derived growth factor-2 (HDGF-2)

Sequence 249 AA;

Query Match 45.0%; Score 1291; DB 6; Length 249;
 Best Local Similarity 99.6%; Pred. No. 3e-110;
 Matches 243; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAASVLRGLDVLVWGLGRYPWPFGKIVNPPKOLKPRGKKCFVVKFGTDEHAWIKVEQL 60
 Db 1 MAASVLRGLDVLVWGLGRYPWPFGKIVNPPKOLKPRGKKCFVVKFGTDEHAWIKVEQL 60
 QY 61 KPYHAHKEEMIKINKGRFQQAQVDAVEEFLRRAKGDQTSNHSDDKNNRNSSEERSRP 120
 Db 61 KPYHAHKEEMIKINKGRFQQAQVDAVEEFLRRAKGDQTSNHSDDKNNRNSSEERSRP 120

Qy	121	NSGDEKRLSLSEGKVKNNMGEGKKRVSSGSSGSGSKSP	PLKRAEQSPKRGPPPKDEKD	180
Db	121	NSGDEKRLSLSEGKVKNNMGEGKKRVSSGSSGSGSKSP	PLKRAEQSPKRGPPPKDEKD	180
Qy	181	LTIPESSTVKGMMAGFMAAFKMQPTASBPVKDADPHFHHFLLSOTETPAVCYQAITKKLK	240	
Db	181	LTIPESSTVKGMMAGFMAAFKMQPTASBPVKDADPHFHHFLLSOTETPAVCYQAITKKLK	240	
Qy	241	ICEE	244	
Db	241	ICED	244	
RESULT 12				
AAU20553				
ID	AAU20553 standard; protein; 260 AA.			
XX	AAU20553;			
AC	XX			
XX	XX			
DT	04-DEC-2001 (first entry)			
XX	Human secreted protein, Seq ID No 545.			
XX	Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;			
KW	rheumatoid arthritis; antiarteriosclerotic; cardiast; vascular;			
KW	cerebroprotective; thrombolytic; antimicrobial; ophthalmological;			
KW	cytostatic; Alzheimer's disease; Parkinson's disease; human; cancer;			
KW	multiple sclerosis; cancer; hyperproliferative disorder; infection;			
KW	Gaucher's disease; neurological disease; cerebrovascular disorder;			
KW	thrombosis; wound healing.			
XX	Homo sapiens.			
XX	WO200155326-A2.			
PN	02-AUG-2001.			
PD	17-JAN-2001; 2001WO-US001347.			
XX	31-JAN-2000; 2000US-0179065P.			
XX	04-FEB-2000; 2000US-0180628P.			
PR	24-FEB-2000; 2000US-0184664P.			
PR	02-MAR-2000; 2000US-0186350P.			
PR	16-MAR-2000; 2000US-0189874P.			
PR	17-MAR-2000; 2000US-0190076P.			
PR	18-APR-2000; 2000US-0198123P.			
PR	19-MAY-2000; 2000US-0205515P.			
PR	07-JUN-2000; 2000US-0209467P.			
PR	28-JUN-2000; 2000US-0214886P.			
PR	30-JUN-2000; 2000US-0215135P.			
PR	07-JUL-2000; 2000US-0216647P.			
PR	07-JUL-2000; 2000US-0216880P.			
PR	11-JUL-2000; 2000US-0217487P.			
PR	11-JUL-2000; 2000US-0217496P.			
PR	14-JUL-2000; 2000US-0218290P.			
PR	26-JUL-2000; 2000US-0220963P.			
PR	26-JUL-2000; 2000US-0220964P.			
PR	14-AUG-2000; 2000US-0224518P.			
PR	14-AUG-2000; 2000US-0224519P.			
PR	14-AUG-2000; 2000US-0225213P.			
PR	14-AUG-2000; 2000US-0225214P.			
PR	14-AUG-2000; 2000US-0225266P.			
PR	14-AUG-2000; 2000US-0225267P.			
PR	14-AUG-2000; 2000US-0225268P.			
PR	14-AUG-2000; 2000US-0225270P.			
PR	14-AUG-2000; 2000US-0225447P.			
PR	14-AUG-2000; 2000US-0225757P.			
PR	14-AUG-2000; 2000US-0225758P.			
PR	14-AUG-2000; 2000US-0225759P.			
PR	18-AUG-2000; 2000US-0228273P.			
PR	22-AUG-2000; 2000US-0228681P.			
PR	22-AUG-2000; 2000US-0228686P.			

PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 03-DEC-2000; 2000US-0251030P.
 PR 03-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 PR (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI; 2001-451931/48.
 DR N-PSDB; AAS33262.
 XX
 PT New nucleic acids and polypeptides, useful for diagnosing, preventing or
 PT treating medical conditions.
 XX
 PS Claim 11; SEQ ID NO 545; 753pp; English.
 XX
 CC The invention relates to novel isolated nucleic acid molecules (I)
 CC encoding human secreted proteins (II). (I) and (II) are used to prevent,
 CC treat or ameliorate a medical condition in e.g. humans, mice, rabbits,
 CC goats, horses, cats, dogs, chickens or sheep. (I) and (II) may be used in
 CC the prevention, treatment and diagnosis of diseases associated with
 CC inappropriate expression of secreted proteins. (I) and complementary
 CC sequences may also be used as DNA probes in diagnostic assays (e.g.
 CC polymerase chain reactions (PCR)) to detect and quantitate the presence
 CC of similar nucleic acid sequences in samples, and so which patients may
 CC be in need of restorative therapy. (II) may also be used as antigens in
 CC the production of antibodies and in assays to identify modulators
 CC (agonists and antagonists) of the expression and activity of the secreted
 CC proteins. The anti-(II) antibodies and antagonists may also be used to
 CC down regulate expression and activity of (II). The anti-(II) antibodies
 CC may also be used as diagnostic agents for detecting the presence of (II)
 CC in samples (e.g. by enzyme linked immunosorbent assay (ELISA)). The
 CC disorders include for example: immune/autoimmune diseases (e.g. HIV
 CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis
 CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.
 CC melanomas, neoplasms of the breast or liver, Sezary syndrome and
 CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,
 CC Parkinson's disease and Charcot-Marie-Tooth disease), cardio-/
 CC cerebrovascular disorders (e.g. cardiac arrest, tachycardia, angina and
 CC thrombosis), infections caused by bacteria, viruses and fungi and ocular
 CC disorders (e.g. corneal infections). (I) and (II), agonists, antagonists
 CC and antibodies can also be used to promote wound healing, maintain organs
 CC before transplantation, and support cell culture of primary tissues.
 CC
 CC Query Match 45.0%; Score 1291; DB 4; Length 260;
 CC Best Local Similarity 99.6%; Pred. No. 3.2e-110;

Matches 243; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MAAVSLRLGLVWGKLGRIYPPWPKI VNP PKLKKPRGKKCFYVFFGTEDHAWIKVEQL 60
 Db 12 MAAVSLRLGLVWGKLGRIYPPWPKI VNP PKLKKPRGKKCFYVFFGTEDHAWIKVEQL 71
 Qy 61 KPYHAHKEEMIKNKGRFOQAVDAVEEFLRRRAKGDQTSNHSDDKNRNSSEERSRP 120
 Db 72 KPYHAHKEEMIKNKGRFOQAVDAVEEFLRRRAKGDQTSNHSDDKNRNSSEERSRP 131
 Qy 121 NSGDEKELSLSEGKVKNNMGEKKRVSSGSSERGSKSLKRAQEQSPRKRGRPPKDEKD 180
 Db 132 NSGDEKELSLSEGKVKNNMGEKKRVSSGSSERGSKSLKRAQEQSPRKRGRPPKDEKD 191
 Qy 181 LTIPESSTVKGMMAGPMAAFKWPQTAPEVKDADPHPHFLLSQTEKPAVCYQAITKKLK 240
 Db 192 LTIPESSTVKGMMAGPMAAFKWPQTAPEVKDADPHPHFLLSQTEKPAVCYQAITKKLK 251
 Qy 241 ICEE 244
 Db 252 ICED 255
 RESULT 13
 AAU21747
 ID AAU21747 standard; protein; 260 AA.
 XX AAU21747;
 AC AAU21747;
 XX
 DT 04-DEC-2001 (first entry)
 XX
 DE Novel human neoplastic disease associated polypeptide #180.
 XX
 KW Human, neoplastic disease associated polypeptide; cancer;
 KW hyperproliferative disorder; neural disorder; immune system disorder;
 KW muscular disorder; reproductive disorder; gastrointestinal disorder;
 KW pulmonary disorder; cardiovascular disorder; renal disorder;
 KW neuroprotective; cytostatic; anti inflammatory; vasotropic.
 XX
 OS Homo sapiens.
 XX
 PN WO200155163-A1.
 XX
 PD 02-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US001358.
 XX
 PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 11-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 26-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.

PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246609P.
PR	14-AUG-2000;	2000US-0225757P.	PR	08-NOV-2000;	2000US-0246610P.
PR	14-AUG-2000;	2000US-0225758P.	PR	08-NOV-2000;	2000US-0246611P.
PR	14-AUG-2000;	2000US-0225759P.	PR	08-NOV-2000;	2000US-0246613P.
PR	18-AUG-2000;	2000US-0226279P.	PR	17-NOV-2000;	2000US-0249207P.
PR	22-AUG-2000;	2000US-0226681P.	PR	17-NOV-2000;	2000US-0249208P.
PR	22-AUG-2000;	2000US-0226688P.	PR	17-NOV-2000;	2000US-0249209P.
PR	22-AUG-2000;	2000US-0227182P.	PR	17-NOV-2000;	2000US-0249210P.
PR	23-AUG-2000;	2000US-0227009P.	PR	17-NOV-2000;	2000US-0249211P.
PR	30-AUG-2000;	2000US-0228924P.	PR	17-NOV-2000;	2000US-0249211P.
PR	01-SEP-2000;	2000US-0228287P.	PR	17-NOV-2000;	2000US-0249212P.
PR	01-SEP-2000;	2000US-0229343P.	PR	17-NOV-2000;	2000US-0249213P.
PR	01-SEP-2000;	2000US-0229344P.	PR	17-NOV-2000;	2000US-0249214P.
PR	01-SEP-2000;	2000US-0229345P.	PR	17-NOV-2000;	2000US-0249215P.
PR	03-SEP-2000;	2000US-0229509P.	PR	17-NOV-2000;	2000US-0249216P.
PR	05-SEP-2000;	2000US-0229513P.	PR	17-NOV-2000;	2000US-0249216P.
PR	06-SEP-2000;	2000US-0230437P.	PR	17-NOV-2000;	2000US-0249217P.
PR	06-SEP-2000;	2000US-0230438P.	PR	17-NOV-2000;	2000US-0249218P.
PR	08-SEP-2000;	2000US-0231242P.	PR	17-NOV-2000;	2000US-0249244P.
PR	08-SEP-2000;	2000US-0231243P.	PR	17-NOV-2000;	2000US-0249245P.
PR	08-SEP-2000;	2000US-0231244P.	PR	17-NOV-2000;	2000US-0249264P.
PR	08-SEP-2000;	2000US-0231413P.	PR	17-NOV-2000;	2000US-0249265P.
PR	08-SEP-2000;	2000US-0231414P.	PR	17-NOV-2000;	2000US-0249297P.
PR	08-SEP-2000;	2000US-0232080P.	PR	17-NOV-2000;	2000US-0249299P.
PR	08-SEP-2000;	2000US-0232081P.	PR	17-NOV-2000;	2000US-0249300P.
PR	12-SEP-2000;	2000US-0231968P.	PR	01-DEC-2000;	2000US-0250160P.
PR	14-SEP-2000;	2000US-0232397P.	PR	01-DEC-2000;	2000US-0250391P.
PR	14-SEP-2000;	2000US-0232398P.	PR	05-DEC-2000;	2000US-0251030P.
PR	14-SEP-2000;	2000US-0232399P.	PR	05-DEC-2000;	2000US-0251988P.
PR	14-SEP-2000;	2000US-0232400P.	PR	05-DEC-2000;	2000US-0256719P.
PR	14-SEP-2000;	2000US-0232401P.	PR	06-DEC-2000;	2000US-0251479P.
PR	14-SEP-2000;	2000US-0233063P.	PR	08-DEC-2000;	2000US-0251856P.
PR	14-SEP-2000;	2000US-0233064P.	PR	08-DEC-2000;	2000US-0251868P.
PR	14-SEP-2000;	2000US-0233065P.	PR	08-DEC-2000;	2000US-0251869P.
PR	21-SEP-2000;	2000US-0234223P.	PR	08-DEC-2000;	2000US-0251989P.
PR	21-SEP-2000;	2000US-0234274P.	PR	08-DEC-2000;	2000US-0251990P.
PR	25-SEP-2000;	2000US-0234997P.	PR	11-DEC-2000;	2000US-0254097P.
PR	25-SEP-2000;	2000US-0234998P.	PR	05-JAN-2001;	2001US-0259678P.
PR	26-SEP-2000;	2000US-0235484P.	XX	(HUMA-) HUMAN GENOME SCI INC.	
PR	27-SEP-2000;	2000US-0235834P.	XX	Rosen CA, Barash SC, Ruben SM;	
PR	27-SEP-2000;	2000US-0235836P.	XX	WPI; 2001-465558/50.	
PR	29-SEP-2000;	2000US-0236327P.	XX	N-PSDD; AAS34946.	
PR	29-SEP-2000;	2000US-0236367P.	DR	Novel polypeptides and polynucleotides useful as diagnostic reagents to	
PR	29-SEP-2000;	2000US-0236368P.	DR	diagnose diseases or disorders associated with aberrant expression or	
PR	29-SEP-2000;	2000US-0236369P.	XX	activity of polypeptides, and for treating cancers, rheumatoid arthritis.	
PR	29-SEP-2000;	2000US-0236370P.	PT	Claim 11; SEQ ID NO 474; 687pp; English.	
PR	29-SEP-2000;	2000US-0236370P.	PT	The present invention relates to the isolation of novel human neoplastic	
PR	02-OCT-2000;	200			

Matches 243; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAASVRLGDLVWGKLGRIYPPWPGKIVNPPKDLKKPRGKCKFFVKFFGTEDHAWLKVEQL 60
Db |||||
Qy 12 MAASVRLGDLVWGKLGRIYPPWPGKIVNPPKDLKKPRGKCKFFVKFFGTEDHAWLKVEQL 71
Db |||||
Qy 61 KPYHAKEEMIKKNGKRFQAVDAVEFLPRAKGDQTSNHSDDKNNRNSSEERSRP 120
Db |||||
Qy 72 KPYHAKEEMIKKNGKRFQAVDAVEFLPRAKGDQTSNHSDDKNNRNSSEERSRP 131
Db |||||
Qy 121 NSGDEKRLSLSEGVKKNMGEGKKRVSSGSGSKSPKRAQFOSPRKGRPPKDEK 180
Db |||||
Qy 132 NSGDEKRLSLSEGVKKNMGEGKKRVSSGSGSKSPKRAQFOSPRKGRPPKDEK 191
Db |||||
Qy 181 LTIPESTVKGMMAGPMAAFKMQPTASEPVKADPHFHHFLLSQTETKPAVCYQAITKKL 240
Db |||||
Qy 192 LTIPESTVKGMMAGPMAAFKMQPTASEPVKADPHFHHFLLSQTETKPAVCYQAITKKL 251
Db |||||
Qy 241 ICEE 244
Db 252 ICED 255

RESULT 14
ADC46388
ID ADC46388 standard; protein; 260 AA.
XX AC ADC46388;
XX DT 18-DEC-2003 (first entry)
XX DE Human neoplastic disease-associated gene 23 protein #2.
XX KW Neoplastic disease-associated polypeptide; gene therapy;
KW hyperproliferative disease; cancer; autoimmune disorder; diabetes;
KW rheumatoid arthritis; systemic lupus erythematosus; multiple sclerosis;
KW autoimmune thyroiditis; haemolytic anaemia; haematopoietic disorder;
KW haematologic disorder; anaemia; thrombocytopaenia; allergic reaction;
KW asthma; eczema; inflammatory disorder; ischaemia-reperfusion injury;
KW inflammatory bowel disease; Crohn's disease; neurodegenerative disorder;
KW Alzheimer's disease; Parkinson's disease; renal disorder;
KW acute glomerulonephritis; end-stage renal disease;
KW cardiovascular disorder; atherosclerosis; myocarditis;
KW infectious disease; AIDS; cachexia; anorexia; wound healing;
KW epithelial cell proliferation; Human.
XX OS Homo sapiens.
XX PN US2003082758-A1.
XX XX
XX PD 01-MAY-2003.
XX PF 22-MAR-2002; 2002US-00103313.
XX 31-JAN-2000; 2000US-017906SP.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226888P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 06-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234224P.
PR 25-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244826P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.

PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0253678P.
PR 17-JAN-2001; 2001US-00764854.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Ruben SM, Barash SC;
XX WPI; 2003-786918/74.
XX N-PSDB; ADC46104.
XX New isolated human neoplastic disease-associated polypeptides and
XX polynucleotides, useful for diagnosing, preventing, prognosticating or
XX treating medical conditions such as cancer, AIDS, diabetes or Parkinson's
XX disease.
XX Claim 11; SEQ ID NO 474; 302pp; English.
XX The invention relates to one of 238 disclosed human neoplastic disease-
XX associated polypeptides encoded by 171 disclosed cDNA sequences
XX (including their domains, epitopes, full-length proteins, allelic variants
XX or species homologues). Also included are there encoding nucleic acids, a
XX recombinant vector comprising the nucleic acid, a recombinant host cell
XX comprising the nucleic acid (expressing the protein), an isolated
XX antibody that binds specifically to the isolated polypeptide, preventing,
XX treating or ameliorating a medical condition, diagnosing a pathological
XX condition or a susceptibility to a pathological condition in a subject,
XX identifying a binding partner to the polypeptide, identifying an activity
XX in a biological assay, and the gene corresponding to the cDNA sequence.
XX The polypeptides, polynucleotides and antibodies are useful for
XX detecting, preventing, diagnosing, prognosticating, treating or

CC ameliorating medical conditions such as hyperproliferative diseases or
CC cancer, autoimmune disorders (e.g. diabetes, rheumatoid arthritis, psoriasis,
CC systemic lupus erythematosus, multiple sclerosis, autoimmune thyroiditis
CC or haemolytic anaemia), haematopoietic or haematologic disorders (e.g.
CC anaemia or thrombocytopenia), allergic reactions including asthma or
CC eczema, inflammatory disorders (e.g. ischaemia-reperfusion injury,
Query Match 45.0%; Score 1291; DB 7; Length 260;
Best Local Similarity 99.6%; Pred. No. 3.2e-110;
Matches 243; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAAVSLRLGDLVWGKLGRIYPPWPKIYNPKDLKKRGGKCFVKKFGTGDHAWIKVEQL 60
DB 12 MAAVSLRLGDLVWGKLGRIYPPWPKIYNPKDLKKRGGKCFVKKFGTGDHAWIKVEQL 71
QY 61 KPYHAHKEEMIKINKGRFOQAVDAVEEFLRRAGKDKQTSNHNSSDDKNRNSSEERSRP 120
DB 72 KPYHAHKEEMIKINKGRFOQAVDAVEEFLRRAGKDKQTSNHNSSDDKNRNSSEERSRP 131
QY 121 NSGDEKRLSLSEGVKKNMGEGKRVSSGSSRGSKSPLKRAQEQSPKRGPRPKDEKD 180
DB 132 NSGDEKRLSLSEGVKKNMGEGKRVSSGSSRGSKSPLKRAQEQSPKRGPRPKDEKD 191
QY 181 LTIPESSTVKGMMAGPMAAFKWPQTASEPVKADDPHFHFLLSQTEKPAVCYQAITKKLK 240
DB 192 LTIPESSTVKGMMAGPMAAFKWPQTASEPVKADDPHFHFLLSQTEKPAVCYQAITKKLK 251
QY 241 ICBE 244
DB 252 ICED 255
RESULT 15
AAU21590
ID AAU21590 standard; protein; 250 AA.
AC AAU21590;
XX
XX 04-DEC-2001 (first entry)
XX Novel human neoplastic disease associated polypeptide #23.
XX Human; neoplastic disease associated polypeptide; cancer;
XX hyperproliferative disorder; neural disorder; immune system disorder;
XX muscular disorder; reproductive disorder; gastrointestinal disorder;
XX pulmonary disorder; cardiovascular disorder; renal disorder;
XX neuroprotective; cytostatic; anti inflammatory; vasotropic.
XX Homo sapiens.
XX WO200155163-A1.
XX 02-AUG-2001.
XX 17-JAN-2001; 2001WO-US0001358.
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.

26-JUL-2000; 2000US-0220964P.
 14-AUG-2000; 2000US-0224518P.
 14-AUG-2000; 2000US-0224519P.
 14-AUG-2000; 2000US-0225213P.
 14-AUG-2000; 2000US-0225214P.
 14-AUG-2000; 2000US-0225266P.
 14-AUG-2000; 2000US-0225267P.
 14-AUG-2000; 2000US-0225288P.
 14-AUG-2000; 2000US-0225270P.
 14-AUG-2000; 2000US-0225447P.
 14-AUG-2000; 2000US-0225757P.
 14-AUG-2000; 2000US-0225758P.
 14-AUG-2000; 2000US-0225759P.
 18-AUG-2000; 2000US-0226279P.
 22-AUG-2000; 2000US-0226681P.
 22-AUG-2000; 2000US-0226868P.
 22-AUG-2000; 2000US-0227182P.
 23-AUG-2000; 2000US-0227009P.
 30-AUG-2000; 2000US-0228924P.
 01-SEP-2000; 2000US-0229287P.
 01-SEP-2000; 2000US-0229343P.
 01-SEP-2000; 2000US-0229344P.
 01-SEP-2000; 2000US-0229345P.
 05-SEP-2000; 2000US-0229509P.
 05-SEP-2000; 2000US-0229513P.
 06-SEP-2000; 2000US-0230437P.
 06-SEP-2000; 2000US-0230438P.
 08-SEP-2000; 2000US-0231242P.
 08-SEP-2000; 2000US-0231243P.
 08-SEP-2000; 2000US-0231244P.
 08-SEP-2000; 2000US-0231143P.
 08-SEP-2000; 2000US-0231414P.
 08-SEP-2000; 2000US-0232080P.
 08-SEP-2000; 2000US-0232081P.
 12-SEP-2000; 2000US-0231968P.
 14-SEP-2000; 2000US-0232397P.
 14-SEP-2000; 2000US-0232398P.
 14-SEP-2000; 2000US-0232399P.
 14-SEP-2000; 2000US-0232400P.
 14-SEP-2000; 2000US-0232401P.
 14-SEP-2000; 2000US-0233063P.
 14-SEP-2000; 2000US-0233064P.
 14-SEP-2000; 2000US-0233065P.
 21-SEP-2000; 2000US-0234223P.
 25-SEP-2000; 2000US-0234274P.
 25-SEP-2000; 2000US-0234997P.
 26-SEP-2000; 2000US-0234998P.
 26-SEP-2000; 2000US-0235484P.
 27-SEP-2000; 2000US-0235834P.
 27-SEP-2000; 2000US-0235836P.
 29-SEP-2000; 2000US-0236327P.
 29-SEP-2000; 2000US-0236367P.
 29-SEP-2000; 2000US-0236368P.
 29-SEP-2000; 2000US-0236369P.
 29-SEP-2000; 2000US-0236370P.
 02-OCT-2000; 2000US-0236802P.
 02-OCT-2000; 2000US-0237037P.
 02-OCT-2000; 2000US-0237038P.
 02-OCT-2000; 2000US-0237039P.
 02-OCT-2000; 2000US-0237040P.
 13-OCT-2000; 2000US-0239335P.
 13-OCT-2000; 2000US-0239337P.
 20-OCT-2000; 2000US-0240960P.
 20-OCT-2000; 2000US-0241221P.
 20-OCT-2000; 2000US-0241785P.
 20-OCT-2000; 2000US-0241786P.
 20-OCT-2000; 2000US-0241787P.
 20-OCT-2000; 2000US-0241808P.
 20-OCT-2000; 2000US-0241809P.
 01-NOV-2000; 2000US-0244617P.
 08-NOV-2000; 2000US-0244647P.
 08-NOV-2000; 2000US-0246475P.
 08-NOV-2000; 2000US-0246476P.
 08-NOV-2000; 2000US-0246477P.
 08-NOV-2000; 2000US-0246478P.
 08-NOV-2000; 2000US-0246523P.
 08-NOV-2000; 2000US-0246524P.
 08-NOV-2000; 2000US-0246525P.
 08-NOV-2000; 2000US-0246526P.
 08-NOV-2000; 2000US-0246527P.
 08-NOV-2000; 2000US-0246528P.
 08-NOV-2000; 2000US-0246532P.
 08-NOV-2000; 2000US-0246609P.
 08-NOV-2000; 2000US-0246610P.
 08-NOV-2000; 2000US-0246611P.
 08-NOV-2000; 2000US-0246613P.
 17-NOV-2000; 2000US-0249207P.
 17-NOV-2000; 2000US-0249208P.
 17-NOV-2000; 2000US-0249209P.
 17-NOV-2000; 2000US-0249210P.
 17-NOV-2000; 2000US-0249211P.
 17-NOV-2000; 2000US-0249212P.
 17-NOV-2000; 2000US-0249214P.
 17-NOV-2000; 2000US-0249215P.
 17-NOV-2000; 2000US-0249216P.
 17-NOV-2000; 2000US-0249217P.
 17-NOV-2000; 2000US-0249218P.
 17-NOV-2000; 2000US-0249244P.
 17-NOV-2000; 2000US-0249245P.
 17-NOV-2000; 2000US-0249264P.
 17-NOV-2000; 2000US-0249265P.
 17-NOV-2000; 2000US-0249297P.
 17-NOV-2000; 2000US-0249299P.
 17-NOV-2000; 2000US-0249300P.
 01-DEC-2000; 2000US-0250160P.
 01-DEC-2000; 2000US-0250391P.
 05-DEC-2000; 2000US-0251030P.
 05-DEC-2000; 2000US-0251988P.
 05-DEC-2000; 2000US-0256719P.
 06-DEC-2000; 2000US-0251479P.
 08-DEC-2000; 2000US-0251856P.
 08-DEC-2000; 2000US-0251868P.
 08-DEC-2000; 2000US-0251869P.
 08-DEC-2000; 2000US-0251989P.
 08-DEC-2000; 2000US-0251990P.
 11-DEC-2000; 2000US-0254097P.
 05-JAN-2001; 2001US-0259678P.
 (HUMA-) HUMAN GENOME SCI INC.
 Rosen CA, Barash SC, Ruben SM;
 WPI; 2001-465558/50.
 N-PSDB; AAS34789.
 Novel polypeptides and polynucleotides useful as diagnostic reagents to
 diagnose diseases or disorders associated with aberrant expression or
 activity of polypeptides, and for treating cancers, rheumatoid arthritis.
 Claim 11; SEQ ID NO 317; 687pp; English.
 The present invention relates to the isolation of novel human neoplastic
 disease associated polypeptides, and cDNA (AAS34767-AAS35050) and DNA
 sequences encoding for these polypeptides. The sequences of the invention
 are useful in the diagnosis, treatment, prevention and/or prognosis of
 disorders involving neoplastic disease such as hyperproliferative
 disorders (e.g. leukaemia, bone cancer, bladder cancer, brain stem
 glioma, adult liver cancer, childhood cerebellar astrocytoma, or
 Hodgkin's lymphoma). The sequences of the invention may also be useful
 for treating other disorders such as neural disorders, immune system
 disorders, muscular disorders, reproductive disorders, gastrointestinal
 disorders, pulmonary disorders, cardiovascular disorders and renal
 disorders. The polynucleotide sequences of the invention are also useful
 in gene therapy. AU21568-AU21851 represent the novel human neoplastic

Thu Sep 16 09:36:51 2004

us-10-067-482-2.rag

CC disease associated polypeptides of the invention. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 250 AA;

Query Match	44.9%	Score 1286;	DB 4;	Length 250;
Best Local Similarity	99.6%	Pred. No. 8.7e-110;		
Matches 242;	Conservative	1;	Mismatches 0;	Indels 0; Gaps 0;

QY	2	AAVSLRLGDLVWGKLGKRYPPWPKGKIVNPPKDLKKPRGKKCFVVKFFGTEDHAMIKVEQLK	61
Db	3	AAVSLRLGDLVWGKLGKRYPPWPKGKIVNPPKDLKKPRGKKCFVVKFFGTEDHAMIKVEQLK	62
QY	62	PYHAHKEEMIKINKGRKFOQAVDAVEEFLRRAGKQQTSSHNSDDDKRRNSSEERSRPN	121
Db	63	PYHAHKEEMIKINKGRKFOQAVDAVEEFLRRAGKQQTSSHNSDDDKRRNSSEERSRPN	122
QY	122	SGDEKRLSLSEGKVKQNMGEKGRVSSGSSERGSKPLKRAQOSPRKGRPPKDEKDL	181
Db	123	SGDEKRLSLSEGKVKQNMGEKGRVSSGSSERGSKPLKRAQOSPRKGRPPKDEKDL	182
QY	182	TIPESSTVKGMMAGPMAFKWQPTASEPVKDADPHFHHFLLSOTKPAVCYQAITKKIKI	241
Db	183	TIPESSTVKGMMAGPMAFKWQPTASEPVKDADPHFHHFLLSOTKPAVCYQAITKKIKI	242
QY	242	CEE 244	
Db	243	CED 245	

RESULT 16
ADC46231
ID ADC46231 standard; protein; 250 AA.

XX AC ADC46231;

XX DT 18-DEC-2003 (first entry)

XX DE Human neoplastic disease-associated gene 23 protein #1.

XX Neoplastic disease-associated polypeptide; gene therapy;
KW hyperproliferative disease; cancer; autoimmune disorder; diabetes;
KW rheumatoid arthritis; systemic lupus erythematosus; multiple sclerosis;
KW autoimmune thyroiditis; haemolytic anaemia; haematopoietic disorder;
KW haematologic disorder; anaemia; thrombocytopaenia; allergic reaction;
KW asthma; eczema; inflammatory disorder; ischaemia-reperfusion injury;
KW inflammatory bowel disease; Crohn's disease; neurodegenerative disorder;
KW Alzheimer's disease; Parkinson's disease; renal disorder;
KW acute glomerulonephritis; end-stage renal disease;
KW cardiovascular disorder; atherosclerosis; myocarditis;
KW infectious disease; AIDS; cachexia; anorexia; wound healing;
KW epithelial cell proliferation; Human.

XX OS Homo sapiens.

XX PN US2003082758-A1.

XX PD 01-MAY-2003.

XX PF 22-MAR-2002; 2002US-00103313.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR	30-JUN-2000;	2000US-0215135P.
PR	07-JUL-2000;	2000US-0216647P.
PR	07-JUL-2000;	2000US-0216880P.
PR	11-JUL-2000;	2000US-0217487P.
PR	11-JUL-2000;	2000US-0217496P.
PR	14-JUL-2000;	2000US-0218290P.
PR	26-JUL-2000;	2000US-0220963P.
PR	26-JUL-2000;	2000US-0220964P.
PR	14-AUG-2000;	2000US-0224518P.
PR	14-AUG-2000;	2000US-0224519P.
PR	14-AUG-2000;	2000US-0225213P.
PR	14-AUG-2000;	2000US-0225214P.
PR	14-AUG-2000;	2000US-0225266P.
PR	14-AUG-2000;	2000US-0225267P.
PR	14-AUG-2000;	2000US-0225268P.
PR	14-AUG-2000;	2000US-0225270P.
PR	14-AUG-2000;	2000US-0225447P.
PR	14-AUG-2000;	2000US-0225757P.
PR	14-AUG-2000;	2000US-0225758P.
PR	14-AUG-2000;	2000US-0225759P.
PR	18-AUG-2000;	2000US-0226279P.
PR	22-AUG-2000;	2000US-0226681P.
PR	22-AUG-2000;	2000US-0226686P.
PR	22-AUG-2000;	2000US-0227182P.
PR	23-AUG-2000;	2000US-0227009P.
PR	30-AUG-2000;	2000US-0228924P.
PR	01-SEP-2000;	2000US-0229287P.
PR	01-SEP-2000;	2000US-0229343P.
PR	01-SEP-2000;	2000US-0229344P.
PR	01-SEP-2000;	2000US-0229345P.
PR	05-SEP-2000;	2000US-0229509P.
PR	05-SEP-2000;	2000US-0229513P.
PR	06-SEP-2000;	2000US-0230437P.
PR	06-SEP-2000;	2000US-0230438P.
PR	08-SEP-2000;	2000US-0231242P.
PR	08-SEP-2000;	2000US-0231243P.
PR	08-SEP-2000;	2000US-0231244P.
PR	08-SEP-2000;	2000US-0231413P.
PR	08-SEP-2000;	2000US-0231414P.
PR	08-SEP-2000;	2000US-0232080P.
PR	08-SEP-2000;	2000US-0232081P.
PR	12-SEP-2000;	2000US-0231968P.
PR	14-SEP-2000;	2000US-0232397P.
PR	14-SEP-2000;	2000US-0232398P.
PR	14-SEP-2000;	2000US-0232399P.
PR	14-SEP-2000;	2000US-0232400P.
PR	14-SEP-2000;	2000US-0232401P.
PR	14-SEP-2000;	2000US-0233063P.
PR	14-SEP-2000;	2000US-0233064P.
PR	14-SEP-2000;	2000US-0233065P.
PR	21-SEP-2000;	2000US-0234223P.
PR	21-SEP-2000;	2000US-0234274P.
PR	25-SEP-2000;	2000US-0234997P.
PR	25-SEP-2000;	2000US-0234998P.
PR	26-SEP-2000;	2000US-0235484P.
PR	27-SEP-2000;	2000US-0235834P.
PR	27-SEP-2000;	2000US-0235836P.
PR	29-SEP-2000;	2000US-0236327P.
PR	29-SEP-2000;	2000US-0236367P.
PR	29-SEP-2000;	2000US-0236368P.
PR	29-SEP-2000;	2000US-0236369P.
PR	29-SEP-2000;	2000US-0236370P.
PR	02-OCT-2000;	2000US-0236802P.
PR	02-OCT-2000;	2000US-0237037P.
PR	02-OCT-2000;	2000US-0237038P.
PR	02-OCT-2000;	2000US-0237039P.
PR	02-OCT-2000;	2000US-0237040P.
PR	13-OCT-2000;	2000US-0239935P.
PR	13-OCT-2000;	2000US-0239937P.
PR	20-OCT-2000;	2000US-0240960P.
PR	20-OCT-2000;	2000US-0241211P.
PR	20-OCT-2000;	2000US-0241785P.
PR	20-OCT-2000;	2000US-0241786P.

20-OCT-2000; 2000US-0241787P.
 20-OCT-2000; 2000US-0241808P.
 20-OCT-2000; 2000US-0241809P.
 01-NOV-2000; 2000US-0241826P.
 08-NOV-2000; 2000US-0244617P.
 08-NOV-2000; 2000US-0246474P.
 08-NOV-2000; 2000US-0246475P.
 08-NOV-2000; 2000US-0246476P.
 08-NOV-2000; 2000US-0246477P.
 08-NOV-2000; 2000US-0246478P.
 08-NOV-2000; 2000US-0246523P.
 08-NOV-2000; 2000US-0246524P.
 08-NOV-2000; 2000US-0246525P.
 08-NOV-2000; 2000US-0246526P.
 08-NOV-2000; 2000US-0246527P.
 08-NOV-2000; 2000US-0246528P.
 08-NOV-2000; 2000US-0246532P.
 08-NOV-2000; 2000US-0246609P.
 08-NOV-2000; 2000US-0246610P.
 08-NOV-2000; 2000US-0246611P.
 17-NOV-2000; 2000US-0246613P.
 17-NOV-2000; 2000US-0249207P.
 17-NOV-2000; 2000US-0249208P.
 17-NOV-2000; 2000US-0249209P.
 17-NOV-2000; 2000US-0249210P.
 17-NOV-2000; 2000US-0249211P.
 17-NOV-2000; 2000US-0249212P.
 17-NOV-2000; 2000US-0249213P.
 17-NOV-2000; 2000US-0249214P.
 17-NOV-2000; 2000US-0249215P.
 17-NOV-2000; 2000US-0249216P.
 17-NOV-2000; 2000US-0249217P.
 17-NOV-2000; 2000US-0249218P.
 17-NOV-2000; 2000US-0249244P.
 17-NOV-2000; 2000US-0249245P.
 17-NOV-2000; 2000US-0249264P.
 17-NOV-2000; 2000US-0249265P.
 17-NOV-2000; 2000US-0249297P.
 17-NOV-2000; 2000US-0249299P.
 17-NOV-2000; 2000US-0249300P.
 01-DEC-2000; 2000US-0250160P.
 01-DEC-2000; 2000US-0250391P.
 05-DEC-2000; 2000US-0251030P.
 05-DEC-2000; 2000US-0251988P.
 06-DEC-2000; 2000US-0256719P.
 08-DEC-2000; 2000US-0251479P.
 08-DEC-2000; 2000US-0251856P.
 08-DEC-2000; 2000US-0251868P.
 08-DEC-2000; 2000US-0251869P.
 08-DEC-2000; 2000US-0251989P.
 11-DEC-2000; 2000US-0251990P.
 05-JAN-2001; 2000US-0254097P.
 17-JAN-2001; 2001US-00764854.
 (HUMA-) HUMAN GENOME SCI INC.
 Rosen CA, Ruben SM, Barash SC;
 WPI; 2003-786918/74.
 N-PSDB; ADC45947.
 New isolated human neoplastic disease-associated polypeptides and polynucleotides, useful for diagnosing, preventing, prognosticating or treating medical conditions such as cancer, AIDS, diabetes or Parkinson's disease.
 Claim 11; SEQ ID NO 317; 302pp; English.
 The invention relates to one of 238 disclosed human neoplastic disease-associated polypeptides encoded by 171 disclosed cDNA sequences (including their domains, epitopes, full-length proteins, allelic variants or species homologues). Also included are there encoding nucleic acids, a

recombinant vector comprising the nucleic acid, a recombinant host cell comprising the nucleic acid (expressing the protein), an isolated antibody that binds specifically to the isolated polypeptide, preventing, treating or ameliorating a medical condition, diagnosing a pathological condition or a susceptibility to a pathological condition in a subject, identifying a binding partner to the polypeptide, identifying an activity in a biological assay, and the gene corresponding to the cDNA sequence. The polypeptides, polynucleotides and antibodies are useful for detecting, preventing, diagnosing, prognosticating, treating or ameliorating medical conditions such as hyperproliferative diseases or cancer, autoimmune disorders (e.g. diabetes, rheumatoid arthritis, systemic lupus erythematosus, multiple sclerosis, autoimmune thyroiditis or haemolytic anaemia), haematopoietic or haematologic disorders (e.g. anaemia or thrombocytopaenia), allergic reactions including asthma or eczema, inflammatory disorders (e.g. ischaemia-reperfusion injury,

Query Match 44.9%; Score 1286; DB 7; Length 250;
 Best Local Similarity 99.6%; Pred. No. 8.7e-110;
 Matches 242; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 AAVSLRLGLDVLWGKLGRIYPPWPGKIVNPPKDLKKPRGKCKFFVKFGTDEHAWIKVQLK 61
 Db |||||
 QY 3 AAVSLRLGLDVLWGKLGRIYPPWPGKIVNPPKDLKKPRGKCKFFVKFGTDEHAWIKVQLK 62
 Db |||||
 QY 62 PYHAHKEEMIKINKGRFQOAVDAVEFLRAKGDQTSNNSDDDKNRRNSSERSRPN 121
 Db |||||
 QY 63 PYHAHKEEMIKINKGRFQOAVDAVEFLRAKGDQTSNNSDDDKNRRNSSERSRPN 122
 Db |||||
 QY 122 SGDEKRLSLSEGVKKNMGEGKRVSSGSSRGSKSPLKRAEQSPKRGPRPKDKDL 181
 Db |||||
 QY 123 SGDEKRLSLSEGVKKNMGEGKRVSSGSSRGSKSPLKRAEQSPKRGPRPKDKDL 182
 Db |||||
 QY 182 TIPSSTVKGMMAGPMAAFKMQPTASEPVKADPHFHLFSQTEKPAVCYQAITKKLI 241
 Db |||||
 QY 183 TIPSSTVKGMMAGPMAAFKMQPTASEPVKADPHFHLFSQTEKPAVCYQAITKKLI 242
 Db |||||
 QY 242 CEE 244
 Db ||:
 QY 243 CED 245

RESULT 17
 ABB60930
 ID ABB60930 standard; protein; 574 AA.
 AC ABB60930;
 DT 26-MAR-2002 (first entry)
 DE Drosophila melanogaster polypeptide SEQ ID NO 9582.
 DE Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX Drosophila melanogaster.
 XX WO200171042-A2.
 PD 27-SEP-2001.
 XX 23-MAR-2001; 2001WO-US009231.
 PR 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX (PEKE) PE CORP NY.
 XX Venter JC, Adams M, Li PWD, Myers EW;
 XX WPI; 2001-656860/75.
 DR N-PSDB; ABL05033.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more

genes from Drosophila and for elucidating cell signaling and cell-cell interactions.

Disclosure; SEQ ID NO 9582; 21pp + Sequence Listing; English.

The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABU16176-ABU130511), expressed DNA sequences (ABU01840-ABU16175) and the encoded proteins (ABB57737-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 574 AA;

Query Match 26.8%; Score 769.5; DB 4; Length 574;
Best Local Similarity 32.6%; Pred. No. 1.5e-61;
Matches 188; Conservative 108; Mismatches 22; Indels 59; Gaps 13;

QY 19 YPPWPGKIVNPPKDL---KKPRGKKCFVKFFGTEHAWIKVEQLKPYHA-HKEEMIKIN 74
DB 5 FTFWPGMIVDPLDLLSOQRANTKCVF--PFGSRNFAWIEENNIKPEGPWKEELAKVS 62
QY 75 KGKRFQQAQVDAVEFLRAKAGKQDTSNNSDDKNRRNSSESRPNRSGDEKRKLISSEG 134
DB 63 KPAARHAMTDIEKYIDDPAEVDEQVKNKSCGAPNHATEADPKIR--DGLDSEISVGEA 120
QY 135 KVKKMGKGRVSGSGSERG-----SKSPL-----KRAEQSPFR----- 169
DB 121 TADGNNGVVAHV--GSPDEGDGLVEINADSSASPVTPAVTTKAGKRTPKAKSVAAAT 178
QY 170 ---KRGPRPKDEKDLTTPESSTVKGMAGPMAAFKQPTAS-BPVKDADP----- 215
DB 179 SVKSTGSAKSAQRRTSAQOSP-----SGPSNAKRGKRDVSGEALQADAEASSTPTGRR 233
QY 216 --HFHFHLLSQTEKPAVCQAITKKLKICEBTGTSTQADSTAVNGSIPTDKKIGFL 273
DB 234 RVETDALLASTAAKRAPNAIALLDPRVTVTRPEAQVIDMSRSNTLADRIDVPESEGTGFL 293
QY 274 GLGMSGIVSNLLKMGHTVTWNRTAEKCDLFIQEGARLGRTPAEVYSTCDITFACVSD 333
DB 294 GLGMSGSTIVKDLITVGHKVVVNNRTIDKQCPFAEGAEVKDTPMDVVEAADVIECCVSD 353
QY 334 PKAAKDLVLGPGSVLQGIKRP--GKCYVDMSTVDADTDTVTELAQVIVSRGRFLFLEAPVSGNQ 391
DB 354 PKGAKDLVFGNGGVLIQ-LKDLANKAYVENSTIDPDTSLDIGBKIQKCNRYLEAQIHGSR 412
QY 392 QLSNDGMLVILAAGDRGLYEDCSSCFQAMGKTSFFLGEVGNAAKMMLIIVNMVQGSFMATI 451
DB 413 QBAEAGMLIILAGGDRSVFECHSCFKTKAKNTFFLGNIGNACKKNLILQTLIGSLVGL 472
QY 452 AEGILAHVTGSOQQLLDILNOGQLASIFLDQKCNILQGNFKPDPLYKYIKQDLRLAI 511
DB 473 AEALALADRFSLNDIIDLFTLSMKSPMLAKGEMAKGDFNPQOPLSHQORLRLVIL 532
QY 512 ALGDVANHPTPMAAAANEVYKRAKALDQSDNDMSAVY 548
DB 533 NMAENLDQSNPEVTISITNEVFKHTKRLGYSEHDSAVF 569

RESULT 18
AAE19927
ID AAE19927 standard; protein; 360 AA.
XX
AC AAE19927;
XX
DT 18-JUN-2002 (first entry)
XX
DE Corn 3-hydroxyisobutyrate dehydrogenase #2.
XX

Branched chain amino acid degradation enzyme; HMG-CoA lyase;
3-hydroxyisobutyrate dehydrogenase; 3-hydroxymethylglutaryl CoA;
isovaleryl-CoA dehydrogenase; transgenic plant; corn; EC 1.1.1.31.

Zea mays.

US6348339-B1.
19-FEB-2002.
29-JUL-1999; 99US-00364230.
31-JUL-1998; 98US-0094990P.
(DUPO) DU PONT DE NEMOURS & CO E I.
Cahoon RE, Hitz WD, Kinney AJ, Rafalski JA;
WPI; 2002-204621/26.
N-PSDB; AAD31748.

Isolated polynucleotide encoding 3-hydroxyisobutyrate dehydrogenase, used to produce transgenic plants having an altered expression of the enzyme.

Claim 1; Col 27-30; 40pp; English.

The invention relates to nucleic acid fragments encoding branched chain amino acid degradation enzymes. Particularly the invention relates to 3-hydroxyisobutyrate dehydrogenase, 3-hydroxymethylglutaryl CoA (HMG-CoA) lyase and isovaleryl-CoA dehydrogenase polypeptides and polynucleotides derived from corn, rice and wheat. Sequences of the invention are used to produce transgenic plants having an altered expression of the enzyme. Polynucleotides of the invention can be used as probes for physical mapping of genomes. The present sequence is corn 3-hydroxyisobutyrate dehydrogenase (EC 1.1.1.31) from crln.pk0191.b10, p0076.cwhan08r and p0118.chsbe01r clones

Sequence 360 AA;

Query Match 25.6%; Score 733; DB 5; Length 360;
Best Local Similarity 52.3%; Pred. No. 1.6e-58;
Matches 148; Conservative 47; Mismatches 88; Indels 0; Gaps 0;

QY 266 TDKTIGTGLGMSGIVSNLLKMGHTVTWNRTAEKCDLFIQEGARLGRTPAEVYSTCD 325
DB 64 SEMEVGLGIGMGKAMATNLLRHGFRVTWNRTAKQCELAALGATVGETPASVVKCR 123
QY 326 ITFACVSDPKAAKDLVLGPGSVLQGIKRPKCYVDMSTVDADTDTVTELAQVIVSRGRFLFLEA 385
DB 124 YTIAMLSDFSALS VVFDKGVLEIQIGSGKGVDMSTVDADTSTKISEAVKQGGAFLEA 183
QY 386 PVSNGNQLSNDGMLVILAAGDRGLYEDCSSCFQAMGKTSFFLGEVGNAAKMMLIIVNMVQ 445
DB 184 PVSNGKPAEDQQLVILAAGDKPLVDGMTIPAFDVLGKKSFFLGEIGNGAKMLVNMVNG 243
QY 446 SPMTAIEGLTAHVTVGSOQQLLDILNOGQLASIFLDQKCNILQGNFKPDPLYKYIK 505
DB 244 SMNLSISEGLCLADKSLSPQLLDVLDLGAITNPMFKLKGPTMLQGSYPAPPLKHQOK 303
QY 506 DLRLATAGDVAHNPTPMAAAANEVYKRAKALDQSDNDMSAVY 548
DB 304 DMRLALGDENAVAMPVSAANEAFKARSILGLGDQDFSAVY 346

RESULT 19
AAE36391
ID AAE36391 standard; protein; 290 AA.
XX
AC AAE36391;
XX
DT 26-JUN-2003 (first entry)
XX
DE Tulip pistil gamma-hydroxybutyrate dehydrogenase homologue.
XX

XX Alpha-methylene-gamma-butyrolactone; glutamate decarboxylase; herbicide;
 KW enzyme; gamma-aminobutyrate aminotransferase; UDP-glucosyltransferase;
 KW gamma-hydroxybutyrate dehydrogenase; tulipalin A; plant.
 XX Tulip pistil.

XX WO2002101013-A2.

XX 19-DEC-2002.

XX 10-JUN-2002; 2002WO-US018230.

XX 08-JUN-2001; 2001US-0297198P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX (PRAB/) PRABHU V.

XX Damude HG, Flint D, Prabhu V, Wang H;

XX WPI; 2003-201331/19.

XX N-PSDB; AAD55024.

XX Novel isolated nucleic acid fragment encoding a tuliposide A synthesizing
 PT protein, useful for creating recombinant organisms that have the ability
 PT to synthesize tulipalin A, tuliposide A or tuliposide A pathway
 PT intermediates.

XX Claim 4; Page 110-111; 71pp; English.

XX The invention relates to genes encoding key enzymes in the biosynthesis
 CC of alpha-methylene-gamma-butyrolactone (tulipalin A). Key enzymes include
 CC glutamate decarboxylase, gamma-aminobutyrate aminotransferase, gamma-
 CC hydroxybutyrate dehydrogenase and UDP-glucosyltransferase. The invention
 CC is useful for producing tulipalin A or tuliposide A or its pathway
 CC intermediates such as alpha-methylenesuccinate semialdehyde, alpha-
 CC methylene-gamma-aminobutyrate or alpha-methylene-gamma-hydroxybutyrate.
 CC Tulipalin A sequences are used to immunise animals to produce polyclonal
 CC or monoclonal antibodies with specificity for them or as targets to
 CC facilitate design and/or identification of inhibitors of those enzymes
 CC that may be useful as herbicides. The present sequence is Tulip pistil
 CC gamma-hydroxybutyrate dehydrogenase homologue

XX Sequence 290 AA;

Query Match 25.1%; Score 720; DB 6; Length 290;

Best Local Similarity 50.7%; Pred. No. 1.8e-57;

Matches 142; Conservative 52; Mismatches 86; Indels 0; Gaps 0;

QY 269 KTGFLGLGMSGIVSNLLKMGHTVTWNRTAEKCDLFTQEGARLGRTPAEVVTCDITF 328

Db 2 EVGFLGLGIMGKAMAVNLLRSGRVTWNRTLSKCNELLEOGASVGETPAAVIKCKYTI 61

QY 329 ACVSDPKAAKDLVLGPGVLOGIRPKCYVDSTVDADVTVELAQVIVSRGRFLPAPVS 388

Db 62 AMLSDPSAALSVPFDKGVLEHMSGKGYDMSVDVNTSSKISEAITQKGGHLEAPVS 121

QY 389 GNQQLSNDGMLVILAAGDRGLYEDSCSCFOAMGKTSFFLGEVGNAAKMMLIVNMVGSFM 448

Db 122 GSKPAEDGQQLIILAAGDKALFEESIPAFDVLGKGSFYLGQVGNANMKMLIVNMVGSFM 181

QY 449 ATIAEGLTIAHVTGOSQOQLLDILNQGLASIFLDQKQNILOQGNPKDPFYLKTIQKDLR 508

Db 182 NALSEGLSLAGKSGLEQKTLDDLDLGAIANPMFKLGKGPAMIQNNHPAPFLKHQCKDMR 241

QY 509 LALALGDVAVNHTPMAAANAENEAFKARSIGLGLDPSAVY 548

Db 242 LALALGDENAVSNPMAAANAENEAFKARSIGLGLDPSAVY 281

RESULT 20

AAE36392

ID AAE36392 standard; protein; 289 AA.

XX

AC AAE36392;

XX 26-JUN-2003 (first entry)

XX

DE A. thaliana gamma-hydroxybutyrate dehydrogenase homologue.

XX

KW Alpha-methylene-gamma-butyrolactone; glutamate decarboxylase; herbicide;
 KW enzyme; gamma-aminobutyrate aminotransferase; UDP-glucosyltransferase;
 KW gamma-hydroxybutyrate dehydrogenase; tulipalin A; plant; mouse-ear cress.

OS Arabidopsis thaliana.

XX WO2002101013-A2.

XX 19-DEC-2002.

XX 10-JUN-2002; 2002WO-US018230.

XX 08-JUN-2001; 2001US-0297198P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX (PRAB/) PRABHU V.

XX Damude HG, Flint D, Prabhu V, Wang H;

XX WPI; 2003-201331/19.

XX N-PSDB; AAD55025.

XX Novel isolated nucleic acid fragment encoding a tuliposide A synthesizing
 PT protein, useful for creating recombinant organisms that have the ability
 PT to synthesize tulipalin A, tuliposide A or tuliposide A pathway
 PT intermediates.

XX Claim 22; Page 112-113; 71pp; English.

XX The invention relates to genes encoding key enzymes in the biosynthesis
 CC of alpha-methylene-gamma-butyrolactone (tulipalin A). Key enzymes include
 CC glutamate decarboxylase, gamma-aminobutyrate aminotransferase, gamma-
 CC hydroxybutyrate dehydrogenase and UDP-glucosyltransferase. The invention
 CC is useful for producing tulipalin A or tuliposide A or its pathway
 CC intermediates such as alpha-methylenesuccinate semialdehyde, alpha-
 CC methylene-gamma-aminobutyrate or alpha-methylene-gamma-hydroxybutyrate.
 CC Tulipalin A sequences are used to immunise animals to produce polyclonal
 CC or monoclonal antibodies with specificity for them or as targets to
 CC facilitate design and/or identification of inhibitors of those enzymes
 CC that may be useful as herbicides. The present sequence is mouse-ear cress
 CC (Arabidopsis thaliana) gamma-hydroxybutyrate dehydrogenase homologue

XX Sequence 289 AA;

Query Match 24.2%; Score 695; DB 6; Length 289;

Best Local Similarity 49.6%; Pred. No. 3.7e-55;

Matches 140; Conservative 49; Mismatches 93; Indels 0; Gaps 0;

QY 269 KTGFLGLGMSGIVSNLLKMGHTVTWNRTAEKCDLFTQEGARLGRTPAEVVTCDITF 328

Db 2 EVGFLGLGIMGKAMSNLLKMGKVTWNRTLSKCDLVEHGASVCEPAEVIKCKYTI 61

QY 329 ACVSDPKAAKDLVLGPGVLOGIRPKCYVDSTVDADVTVELAQVIVSRGRFLPAPVS 388

Db 62 AMLSDPSAALSVPFDKGVLEQICGKGYDMSVDVNTSSKISEAITQKGGHLEAPVS 121

QY 389 GNQQLSNDGMLVILAAGDRGLYEDSCSCFOAMGKTSFFLGEVGNAAKMMLIVNMVGSFM 448

Db 122 GSKPAEDGQQLIILAAGDKALFEESIPAFDVLGKGSFYLGQVGNANMKMLIVNMVGSFM 181

QY 449 ATIAEGLTIAHVTGOSQOQLLDILNQGLASIFLDQKQNILOQGNPKDPFYLKTIQKDLR 508

Db 182 NAFSEGLVLADKSGLSLSDTLDDLDLGAIANPMFKLGKGPAMIQNNHPAPFLKHQCKDMR 241

QY 509 LALALGDVAVNHTPMAAANAENEAFKARSIGLGLDPSAVY 550

Db 242 LALALGDNANVNFVAAAANEAFKARSGLGLDLSAVIEA 283

RESULT 21

AAEL19929 ID AAE19929 standard; protein; 345 AA.

XX AC AAE19929;

XX DT 18-JUN-2002 (first entry)

XX DE Soybean 3-hydroxyisobutyrate dehydrogenase #1.

XX KW Branched chain amino acid degradation enzyme; HMG-CoA lyase;
XX KW 3-hydroxyisobutyrate dehydrogenase; 3-hydroxymethylglutaryl CoA;
XX KW isovaleryl-CoA dehydrogenase; transgenic plant; soybean; EC 1.1.1.31.

XX OS Glycine max.

XX PN US6348339-B1.

XX PD 19-FEB-2002.

XX PF 29-JUL-1999; 99US-00364230.

XX PR 31-JUL-1998; 98US-0094990P.

XX PA (DUPO) DU PONT DE NEMOURS & CO E I.

XX PI Cahoon RE, Hitz WD, Kinney AJ, Rafalski JA;

XX DR WPI; 2002-204621/26.

XX DR N-PSDB; RAD31750.

XX PT Isolated polynucleotide encoding 3-hydroxyisobutyrate dehydrogenase, used
XX FT to produce transgenic plants having an altered expression of the enzyme.

XX PS Claim 1; Col 33-36; 40pp; English.

XX CC The invention relates to nucleic acid fragments encoding branched chain
XX CC amino acid degradation enzymes. Particularly the invention relates to 3-
XX CC hydroxyisobutyrate dehydrogenase, 3-hydroxymethylglutaryl CoA (HMG-CoA)
XX CC lyase and isovaleryl-CoA dehydrogenase polypeptides and polynucleotides
XX CC derived from corn, rice and wheat. Sequences of the invention are used to
XX CC produce transgenic plants having an altered expression of the enzyme.
XX CC Polynucleotides of the invention can be used as probes for physical
XX CC mapping of genomes. The present sequence is soybean 3-hydroxyisobutyrate
XX CC dehydrogenase (EC 1.1.1.31) from ses2w.pk0002.e5 clone

XX SQ Sequence 345 AA;

Query Match 23.8%; Score 682; DB 5; Length 345;

Best Local Similarity 47.1%; Pred. No. 7.7e-54;

Matches 136; Conservative 54; Mismatches 99; Indels 0; Gaps 0;

QY 262 STPTDKKIGFGLGMLSGGIVSNLLKMGHTTVNVRTAEKCDLFTQEGARLGRTPAEVY 321

DB 46 AVTEPPARIGFGLGMLSGGIVSNLLKMGHTTVNVRTAEKCDLFTQEGARLGRTPAEVY 105

QY 322 STCDITFACVSPKAAKDLVLGPGSGVLOGIRPKCYVDMSTVDADTVTELAVIVSRGGR 381

DB 106 ASCDVTFAMLADPQSNVAVACGKHGAANGPGKGVVDVSTVDGDTSLKINGHMKSTGAL 165

QY 382 FLEAPVSGNQOISNDGMLVILAGDRGLVEDSCSCFQAMGKTSFFLGEVGNAAKMLIVN 441

DB 166 FLEAPVSGSKPAEDQQLIFLTAGDKNLYEAVGSLDDINGKSKFYLDGVNGAAMKLIVN 225

QY 442 MVQSGFWATAEGLTLAHVTGQSQOQLDILNQGQSLASIFLDQKCNILQGNFKPDPFLK 501

DB 226 MINGSMWAFSEGLLSEKVDLPDVLVQVVSQALSAIPMYSTKGFMSIQSLYPTAFPLK 285

QY 502 YIQDLRLALGLDANVNHPTPMAAAANEVYKRAKALDQSDNDMSAVYRA 550

Db 286 HQQKRLRLALGLAESVSVQPTPIASAANELYKVAKNSGLSDQDFSAVIEA 334

RESULT 22

AAAB19491 ID AAB19491 standard; protein; 289 AA.

XX AC AAB19491;

XX DT 06-MAR-2001 (first entry)

XX DE Amino acid sequence of gamma-hydroxy butyrate dehydrogenase (GHBHDH).

XX KW Gamma-aminobutyric acid; GABA; GABA transaminase; GABA-T;
XX KW gamma-hydroxy butyrate dehydrogenase; GHBHDH; succinic semialdehyde;
XX KW gamma hydroxybutyrate.

XX OS Arabidopsis thaliana.

XX PN Key Location/Qualifiers

XX FT Misc-difference 1..100 /note= "these residues are encoded by nucleotides 1-300"

XX FT WO200061763-A2.

XX PD 19-OCT-2000.

XX PF 06-APR-2000; 2000WO-CA000378.

XX PR 09-APR-1999; 99US-0128518P.

XX PA (UYGU-) UNIV GUELPH.

XX PI Shelp BJ, Breitzkreuz KE, Van Cauwenberghe OR;

XX DR WPI; 2000-679492/66.

XX DR N-PSDB; AAC62027.

XX PT Novel proteins related to gamma aminobutyric acid (GABA) metabolism, GABA
XX FT -transaminase and gamma hydroxy butyrate dehydrogenase useful for
XX FT screening and isolation of homologous genes from other organisms.

XX PS Claim 18; Fig 5; 48pp; English.

XX CC The present sequence represents a plant gamma-hydroxy butyrate
XX CC dehydrogenase (GHBHDH). The specification also describes a plant gamma-
XX CC aminobutyric acid (GABA) transaminase (GABA-T). The proteins are involved
XX CC in the metabolism of GABA in plants. GABA-T is used to transaminase GABA
XX CC to form succinic semialdehyde, which may be converted to gamma
XX CC hydroxybutyrate by GHBHDH. The polynucleotides are used to produce GABA-T
XX CC and GHBHDH recombinantly. The proteins are useful for screening and
XX CC isolation of homologous genes from other organisms

XX SQ Sequence 289 AA;

Query Match 23.7%; Score 679; DB 3; Length 289;

Best Local Similarity 48.6%; Pred. No. 1.1e-53;

Matches 137; Conservative 49; Mismatches 96; Indels 0; Gaps 0;

QY 269 KTGFLGLGMLSGGIVSNLLKMGHTTVNVRTAEKCDLFTQEGARLGRTPAEVYSTCDITF 328

DB 2 EYGFGLGIMGKAMGNLLKNGFKVTVMNRTLSKCDLVEHGASVCESPAEVIKKCKYTI 61

QY 329 ACVSDPKAAKDLVLGPGSGVLOGIRPKCYVDMSTVDADTVTELAVIVSRGGRFLEAPVS 388

DB 62 AMLSOPCAALSVDKGGVLEQICEGKYIDMSTVDAETSLKINEAITGKGRFVEGPPVS 121

QY 389 GNOQISNDGMLVILAGDRGLVEDSCSCFQAMGKTSFFLGEVGNAAKMLIVNMVQSGSEM 448

DB 122 GSKKPAEDQQLIFLTAGDKNLYEAVGSLDDINGKSKFYLDGVNGAAMKLIVNMVQSGSEM 181

QY 449 ATTAEGTLAHVTGQSQOQLDILNQGQSLASIFLDQKCNILQGNFKPDPFLKYIQDLR 508

PR	13-AUG-1999;	99US-0148565P.	Db	51	TTPSTRDELGTVS-----IGFLGIMSGFMAQNLIKAGCDVTVMN	91
PR	13-AUG-1999;	99US-0148684P.	Qy	298	RTAEKCDLFIQEGARLGRTPAEVVSTCDITFACVSDPKAAKDLVLGPGSVLQGIROPKCY	357
PR	16-AUG-1999;	99US-01491368P.	Db	92	RTSKCDPLVLGAKYKSSPEEVTACDLTFAMLADESAIDVACGKNGAIFGISSGKY	151
PR	17-AUG-1999;	99US-0149175P.	Qy	358	VDMSITVDADTVTELAQVIVSRGGRFLEAPVSGNQQLSNDGMLVILAAGDRGLHEDCSCEF	417
PR	18-AUG-1999;	99US-0149426P.	Db	152	VDVSTVDVASSILISKIKOTGALFLEAPVSGSKKPAEDGQLIFLTAGDKFELYEKAPFL	211
PR	20-AUG-1999;	99US-0149722P.	Qy	418	QAMGKTSFFLGEVNAKMMLIYNNVOGSFWATTAEGTLTAHVTGQSQOQLDILINOGOL	477
PR	20-AUG-1999;	99US-0149929P.	Db	212	DIMGSKFYLGEVNGAAMKLVNMMINGSMWASFAEGILLSSQKVLDPNVLVEVVSQAI	271
PR	20-AUG-1999;	99US-0149930P.	Qy	478	ASIFLDOKCONILOKFKPDYLYKYIQKDLRLALDALGDVANNHPTPMAAANVEVYKAKAL	537
PR	23-AUG-1999;	99US-0150566P.	Db	272	NAPMYSLKGFMSMKSVVPTAFPLKHQOKDMELALGLAESVSQSTPIAAAANELYKVAKY	331
PR	26-AUG-1999;	99US-0150884P.	Qy	538	DQSDNDMSAVYRA	550
PR	27-AUG-1999;	99US-0151065P.	Db	332	GLSDEDESAVIEA	344
PR	27-AUG-1999;	99US-0151086P.				
PR	27-AUG-1999;	99US-0151080P.				
PR	30-AUG-1999;	99US-0151303P.				
PR	31-AUG-1999;	99US-0151438P.				
PR	01-SEP-1999;	99US-0151930P.				
PR	07-SEP-1999;	99US-0152363P.				
PR	10-SEP-1999;	99US-0153070P.				
PR	13-SEP-1999;	99US-0153758P.				
PR	15-SEP-1999;	99US-0154018P.				
PR	16-SEP-1999;	99US-0154039P.				
PR	20-SEP-1999;	99US-0154779P.				
PR	22-SEP-1999;	99US-0155139P.				
PR	23-SEP-1999;	99US-0155486P.				
PR	24-SEP-1999;	99US-0155659P.				
PR	28-SEP-1999;	99US-0156458P.				
PR	29-SEP-1999;	99US-0156596P.				
PR	04-OCT-1999;	99US-0157117P.				
PR	05-OCT-1999;	99US-0157753P.				
PR	06-OCT-1999;	99US-0157865P.				
PR	07-OCT-1999;	99US-0158029P.				
PR	08-OCT-1999;	99US-0158232P.				
PR	12-OCT-1999;	99US-0158369P.				
PR	13-OCT-1999;	99US-0159293P.				
PR	13-OCT-1999;	99US-0159294P.				
PR	13-OCT-1999;	99US-0159295P.				
PR	13-OCT-1999;	99US-0159329P.				
PR	14-OCT-1999;	99US-0159329P.				
PR	14-OCT-1999;	99US-0159331P.				
PR	14-OCT-1999;	99US-0159637P.				
PR	14-OCT-1999;	99US-0159638P.				
PR	14-OCT-1999;	99US-0159848P.				
PR	18-OCT-1999;	99US-0160741P.				
PR	21-OCT-1999;	99US-0160767P.				
PR	21-OCT-1999;	99US-0160768P.				
PR	21-OCT-1999;	99US-0160770P.				
PR	21-OCT-1999;	99US-0160814P.				
PR	21-OCT-1999;	99US-0160815P.				
PR	22-OCT-1999;	99US-0160980P.				
PR	22-OCT-1999;	99US-0160981P.				
PR	22-OCT-1999;	99US-0160989P.				
PR	25-OCT-1999;	99US-0161404P.				
PR	25-OCT-1999;	99US-0161405P.				
PR	25-OCT-1999;	99US-0161406P.				
PR	26-OCT-1999;	99US-0161359P.				
PR	26-OCT-1999;	99US-0161360P.				
PR	26-OCT-1999;	99US-0161361P.				
PR	28-OCT-1999;	99US-0161920P.				
PR	28-OCT-1999;	99US-0161922P.				
PR	28-OCT-1999;	99US-0161933P.				
PR	29-OCT-1999;	99US-0162142P.				
Query Match 23.3%; Score 667.5; DB 3; Length 353;						
Best Local Similarity 39.1%; Pred. No. 1.7e-52;						
Matches 146; Conservative 65; Mismatches 121; Indels 41; Gaps 5;						
Qy	186	SSTVKGM---MAGPMAAFKWOPTASEPVKDADPHFHHLLSQTEKPAVC-----YQAITK	237			
Db	5	SSSSKAMALCSICPRILPRFRPKPISE-----FL-----SKPQICLAYRVYSLSLQ	50			
Qy	238	KLKICEETGTSIQADSTAVNGSIPTDKKIGFLGLGIMSGSIVSNLLKMGHTVTVNM	297			


```

RESULT 25
AAG10425
ID  AAG10425 standard; protein; 285 AA.
XX
XX  AC
XX  AAG10425;
XX
XX  17-OCT-2000 (first entry)
XX
XX  Arabidopsis thaliana protein fragment SEQ ID NO: 8741.
DE
XX
XX  Protein identification; signal transduction pathway; metabolic pathway;
KW  hybridisation assay; genetic mapping; gene expression control; promoter;
KW  termination sequence.
XX
XX  Arabidopsis thaliana.
OS
XX  EP1033405-A2.
PN
XX  06-SEP-2000.
PD
XX
XX
XX  25-FEB-2000; 2000EP-00301439.
PF
XX
XX  25-FEB-1999; 99US-0121825P.
PR  05-MAR-1999; 99US-0123180P.
PR  09-MAR-1999; 99US-0123548P.
PR  23-MAR-1999; 99US-0125788P.
PR  25-MAR-1999; 99US-0126264P.
PR  29-MAR-1999; 99US-0126785P.
PR  01-APR-1999; 99US-0127462P.
PR  06-APR-1999; 99US-0128234P.
PR  08-APR-1999; 99US-0128714P.
PR  16-APR-1999; 99US-0129845P.
PR  19-APR-1999; 99US-0130077P.
PR  21-APR-1999; 99US-0130449P.
PR  23-APR-1999; 99US-0130510P.
PR  23-APR-1999; 99US-0130891P.

```

```

PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147433P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 16-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149802P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150366P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0156559P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.

PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 21.9%; Score 629; DB 3; Length 285;
Best Local Similarity 45.3%; Pred. No. 4.4e-49;
Matches 125; Conservative 56; Mismatches 95; Indels 0; Gaps 0;

QY 275 LGLMGSGIVSNLLKMGHTVTWRTAEKCDLFTQEGARLGRTPAEVTVSTCDITFACVSDP 334
DB 1 MGIMGSPMAQNLKAGCDVTWRTKSKCDPLVGLGAKYKSSPEEVATCDLTFAMLADP 60

QY 335 KAAKDLVLGPGVLOQIRPGKCYVDMSTVDADTVTELAQVIVSRGGRFLEAPVSGNQOLS 394
DB 61 ESALDVACGKNGAIFGLSSGKGVVDVSTVVASILISKQIKDTGALFLEAPVSGSKKPA 120

QY 395 NDMVLILAAAGDRGLYEDCSSCFQAMKTSFFLGEVGNAAKMLIVNMVQGSFMATTIAG 454
DB 121 EDGQLIFLTAGDKPLYEKAAPFLDIMGSKFYLGVCNGAAMKLVNWMINGSWWSFAEG 180

QY 455 LTLAHTVGSQQOFTLLDLNQGLASFLDOKCONIQGNFKPDFYLYKIOKDLRLAIALG 514
DB 181 ILLSQKVLDPNVLVEVWSQAINAPMYSILKGPSMKSVYPTAFPLKHQOKDMRLAIGLA 240

QY 515 DAVNHPTMAAAANVYKRAKALDQSDNDMSAVYRA 550
DB 241 ESVSQSTPIAAAANELYKVAKSYGLSDEDFSAVIEA 276

RESULT 26
AAE19926
ID AAE19926 standard; protein; 247 AA.
XX AAE19926;
XX
XX
DT 18-JUN-2002 (first entry)
XX
DE Corn 3-hydroxyisobutyrate dehydrogenase #1.
XX
XX Branched chain amino acid degradation enzyme; HMG-CoA lyase;
XX 3-hydroxyisobutyrate dehydrogenase; 3-hydroxymethylglutaryl CoA;
XX isovaleryl-CoA dehydrogenase; transgenic plant; corn; EC 1.1.1.31.
XX
XX Zea mays.
XX
XX Key Location/Qualifiers
FH Misc-difference 244 /label= Unknown
FT /note= "Encoded by NCT"
FT Misc-difference 247 /label= Unknown
FT /note= "Encoded by NTG"
XX
XX US6348339-B1.
XX
XX 19-FEB-2002.

```

Thu Sep 16 09:36:51 2004

us-10-067-482-2.ra.g

```
XX 29-JUL-1999; 99US-00364230.
XX PF
XX 31-JUL-1998; 98US-0094990P.
XX PR
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
XX PA
XX Cahoon RE, Hitz WD, Kinney AJ, Rafalski JA;
XX WPI; 2002-204621/26.
XX DR N-PSDB; AAD31747.
XX DR
XX Isolated polynucleotide encoding 3-hydroxyisobutyrate dehydrogenase, used
XX PT to produce transgenic plants having an altered expression of the enzyme.
XX PS Disclosure; Col 23-26; 40pp; English.
XX CC The invention relates to nucleic acid fragments encoding branched chain
XX CC amino acid degradation enzymes. Particularly the invention relates to 3-
XX CC hydroxyisobutyrate dehydrogenase, 3-hydroxymethylglutaryl CoA (HMG-CoA)
XX CC lyase and isovaleryl-CoA dehydrogenase polypeptides and polynucleotides
XX CC derived from corn, rice and wheat. Sequences of the invention are used to
XX CC produce transgenic plants having an altered expression of the enzyme.
XX CC Polynucleotides of the invention can be used as probes for physical
XX CC mapping of genomes. The present sequence is corn 3-hydroxyisobutyrate
XX CC dehydrogenase (EC 1.1.1.31) from cemin.pk0138.el, ctain.pk0052.d10 and
XX CC p0037.crwak74r clones
XX SQ Sequence 247 AA;

Query Match 18.8%; Score 537.5; DB 5; Length 247;
Best Local Similarity 45.3%; Pred. No. 5.8e-41;
Matches 106; Conservative 47; Mismatches 80; Indels 1; Gaps 1;

Qy 284 SNLLKMGHTVTVNRTAKCDLFTQEGARLGRTPAEVVTCDITPACVSDPKAAKDLVLG 343
Dy 2 SNLIAGCDTVNRTKSCDPLLSLGAKEYPTPAQVASSCDVTFAMLADPQSAAEVACG 61

Qy 344 PSGLVQGIKPKCYVDMSTVDADVTTELAQVIVSRGGRFLAPVSGNQOLNDGMLVILA 403
Dy 62 SSGAAGLAPGKGYVDVSTVDGATSKLIGERTITSGASFLEAPVSGSKPAEDGLLFLT 121

Qy 404 AGDRGLYEDCSSCFQMGKTSFFELGCVGNAAKMWLVNVOGSEFMATIAEGLTLAHVTGQ 463
Dy 122 AGDESLYKRVAPLLDVMGKSRFYLGDVNGAAMKLVNVMGSMVVSFSEGLLLSEKVL 181

Qy 464 SQOTLLDILNOGLASIFLDQCNILQGNFKPDEFYLYIQKDLRLAIALGDV 517
Dy 182 DPNLLVEVISQGISAPMSESLKGPSMKVKAAYPPAPLKHQKQKLS-SIGLADRV 234

RESULT 27
AAG41222
ID AAG41222 standard; protein; 168 AA.
XX AC
XX AAG41222;
XX DT
XX 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 51261.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX XX
XX EP1033405-A2.
XX PD
XX 06-SEP-2000.
XX PF
XX 25-FEB-2000; 2000EP-00301439.
XX XX
```

```
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 24-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
```

```
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 20-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156459P.
PR 29-SEP-1999; 99US-0156596P.

PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 13.9%; Score 399; DB 3; Length 168;
Best Local Similarity 48.4%; Pred. No. 3.2e-28;
Matches 76; Conservative 29; Mismatches 52; Indels 0; Gaps 0;

QY 269 KTGFLGLMGSGIVSNLLKMGHTVTVMNRTAEKCDLFIQEGARLGRTPAEVVSTCDITF 328
Db 2 EVGFLGLGIMGKAMSNLLKNGFKVTVMNRTLSKCDLVEHGASVCESPAEVIKKCKYTI 61

QY 329 ACVSDPKAAKDLVLGPGSVLGIRPKGCVDMSTVDADTVELAQVIVSRGGRFLEAPVS 388
Db 62 AMLSDPCAALSVMFKGGVLEQICEKGVIDMSTVDAETSLKINEAITGKGRFVEGPVS 121

QY 389 GNQQLSNDGMLVTLAAGDRGLVEDCSSCFQAMGKTSF 425
Db 122 GSKKPAEDGQLILLAAAGKALFEESIPALMSWGRDRF 158

RESULT 28
AAG20956
ID AAG20956 standard; protein; 312 AA.
XX AAG20956;
AC AAG20956;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 23337.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
```


XX PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 16-APR-1999; 99US-0128714P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132487P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135112P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135623P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0139119P.
PR 14-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 26-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.

PR	28-JUN-1999;	99US-0140923P.	PR	10-SEP-1999;	99US-0153070P.
PR	29-JUN-1999;	99US-0140991P.	PR	13-SEP-1999;	99US-0153758P.
PR	30-JUN-1999;	99US-0141287P.	PR	15-SEP-1999;	99US-0154018P.
PR	01-JUL-1999;	99US-0141842P.	PR	16-SEP-1999;	99US-0154039P.
PR	01-JUL-1999;	99US-0142154P.	PR	20-SEP-1999;	99US-0154779P.
PR	02-JUL-1999;	99US-0142055P.	PR	22-SEP-1999;	99US-0155139P.
PR	06-JUL-1999;	99US-0142390P.	PR	23-SEP-1999;	99US-0155486P.
PR	08-JUL-1999;	99US-0142803P.	PR	24-SEP-1999;	99US-0155659P.
PR	09-JUL-1999;	99US-0142920P.	PR	28-SEP-1999;	99US-0156458P.
PR	12-JUL-1999;	99US-0143542P.	PR	29-SEP-1999;	99US-0156596P.
PR	13-JUL-1999;	99US-0143624P.	PR	04-OCT-1999;	99US-0157117P.
PR	14-JUL-1999;	99US-0144005P.	PR	06-OCT-1999;	99US-0157753P.
PR	15-JUL-1999;	99US-0144085P.	PR	06-OCT-1999;	99US-0157865P.
PR	16-JUL-1999;	99US-0144086P.	PR	07-OCT-1999;	99US-0158029P.
PR	16-JUL-1999;	99US-0144325P.	PR	08-OCT-1999;	99US-0158232P.
PR	19-JUL-1999;	99US-0144331P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144334P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159329P.
PR	20-JUL-1999;	99US-0144332P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144844P.	PR	14-OCT-1999;	99US-0159637P.
PR	21-JUL-1999;	99US-0144814P.	PR	14-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-0145086P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145085P.	PR	21-OCT-1999;	99US-0160767P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160814P.
PR	23-JUL-1999;	99US-0145145P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160981P.
PR	26-JUL-1999;	99US-0145276P.	PR	22-OCT-1999;	99US-0160989P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145919P.	PR	25-OCT-1999;	99US-0161406P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161359P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146388P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146389P.	PR	26-OCT-1999;	99US-0161920P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161922P.
PR	04-AUG-1999;	99US-0147204P.	PR	28-OCT-1999;	99US-0161993P.
PR	04-AUG-1999;	99US-0147302P.	PR	29-OCT-1999;	99US-0162142P.
PR	05-AUG-1999;	99US-0147192P.	Query Match 13.4%; Score 384; DB 3; Length 334;		
PR	05-AUG-1999;	99US-0147260P.	Best Local Similarity 32.2%; Pred. No. 2,3e-26;		
PR	06-AUG-1999;	99US-0147303P.	Matches 96; Conservative 48; Mismatches 152; Indels 2; Gaps 2;		
PR	06-AUG-1999;	99US-0147416P.	254 ADSTAVNGSITPTDKKIGLGLGMSGIVSNLLKMGHTVTVNRTAEKCDLFIQEGARL 313		
PR	09-AUG-1999;	99US-0147493P.	24 ASSTISSDIITPSNPKIGMIGTVNMGSMCGHLIKAGYTVTVFNRTISKAQTLIDMGANV 83		
PR	09-AUG-1999;	99US-0147935P.	314 GRTPAEVWSTCDITFACVSDPKAADVLGP-SGVLQGIKPGKCVDMSTVDADTVTELA 372		
PR	10-AUG-1999;	99US-0148171P.	84 ADSPNSVAEQSDVWFTIVGYPSDVRHVLDDPKSGALSGLRQGGVLVDVMTTSPSLAEETA 143		
PR	11-AUG-1999;	99US-0148319P.	373 QVIVSRGGRFLEAPVSGNQQLSNDQMLVILAAAGRLYEDCSCFCQAMGKTSFFLGEVGN 432		
PR	12-AUG-1999;	99US-0148341P.	144 KAASTKNCESIDAPVSGGDLGAKNGKLSIFAGGDETTVKRLDPLFSLMGKVN-FMGTSKG 202		
PR	13-AUG-1999;	99US-0148565P.	433 AAKMMLIVNMVQGSFMATIAEGLTIAHVTGQSQOQLDILNQGQOLASFLDQKCONILQG 492		
PR	13-AUG-1999;	99US-0148684P.	203 GQFALKANQITTIATMLGLIVEGLIYAHKAGLDVKKFLEAISTGAAGSKSIDLYGRILKR 262		
PR	16-AUG-1999;	99US-0149368P.	493 NEKPDPLYKYIOKDLRLALATGDVNNHPTPMAAANEVYKRAKALDQSDNDMSAVYRA 550		
PR	17-AUG-1999;	99US-0149175P.	263 DFDPGFVNHFFVKDLGICLNQCQRMGLALPGLALAAQQLYLSLKHGEGDLGTQALLA 320		
PR	18-AUG-1999;	99US-0149426P.	RESULT 30		
PR	20-AUG-1999;	99US-0149722P.	AAG20954		
PR	20-AUG-1999;	99US-0149723P.	ID AAG20954 standard; protein; 336 AA.		
PR	20-AUG-1999;	99US-0149929P.			
PR	23-AUG-1999;	99US-0149902P.			
PR	23-AUG-1999;	99US-0149930P.			
PR	25-AUG-1999;	99US-0150566P.			
PR	26-AUG-1999;	99US-0150884P.			
PR	27-AUG-1999;	99US-0151065P.			
PR	27-AUG-1999;	99US-0151066P.			
PR	27-AUG-1999;	99US-0151080P.			
PR	30-AUG-1999;	99US-0151303P.			
PR	31-AUG-1999;	99US-0151438P.			
PR	01-SEP-1999;	99US-0151930P.			
PR	07-SEP-1999;	99US-0152363P.			

PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-01511303P.
PR	30-AUG-1999;	99US-0151438P.
PR	31-AUG-1999;	99US-0151930P.
PR	01-SEP-1999;	99US-0152363P.
PR	07-SEP-1999;	99US-0153070P.
PR	10-SEP-1999;	99US-0153758P.
PR	13-SEP-1999;	99US-0154018P.
PR	15-SEP-1999;	99US-0154039P.
PR	16-SEP-1999;	99US-0154779P.
PR	20-SEP-1999;	99US-0155139P.
PR	22-SEP-1999;	99US-0155486P.
PR	23-SEP-1999;	99US-0155659P.
PR	24-SEP-1999;	99US-0156458P.
PR	28-SEP-1999;	99US-0156596P.
PR	29-SEP-1999;	99US-0157117P.
PR	04-OCT-1999;	99US-0157753P.
PR	05-OCT-1999;	99US-0157865P.
PR	06-OCT-1999;	99US-0158029P.
PR	07-OCT-1999;	99US-0158232P.
PR	08-OCT-1999;	99US-0158369P.
PR	12-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
Query Match 13.4%; Score 384; DB 3; Length 336;		
Best Local Similarity 32.2%; Pred No. 2.3e-26;		
Matches 96; Conservative 48; Mismatches 152; Indels 2; Gaps 2;		
QY	254	ADSTAVNGSIPTDKKIGLGLGMLSGSIVNLLKXGHTVTVMNRTAEKCDLFIQEGARL 313
DB	26	ASSTISSDIIIPSNFKIGWIGVNGRSMCGHLIKAGYTVTVFNRTISKQATLIDMGANV 85
QY	314	GRTPAEVVTCDITFACVSDPKAAKDLVLP-SGVLOGIRPGKCYVDMSTVDADIVTELA 372
DB	86	ADSPNSVABQSDVVFIVGYPSDVRHVLDDPKSGALSGRLQGGLVDMVTTSFSLAEIFA 145
QY	373	QVIVSRGGRFLEAPVSGNQQLSNDGMLVTLAAGDRGLVEDCSSCFQAMGKTSFFLGEVGN 432
DB	146	KAASFNCFSIDAPVSGGDLGAKNGKLSIFAGGDETVTKRLDPLFSLMGKVN-FMGTSK 204
QY	433	AAKMMILVNNVQGSFMATTAEGTLTAHVTVQSQQTLLDILNOGOLASFIDOKCQNILOQ 492
DB	205	GQFAKLANQITIASTMLGLVEGLIVAHKAGLDVKKFLEAISTGAAGSKSIDLYGDRILKR 264

QY 493 NFKPDPFLKYIQKDLRLALALGDAVNHPTPMAAAANEVYKAKALDQSDNDMGAVYRA 550
DB 265 DFDPGFVNVHFVKDLGICLNECCORMGLALPGLAQAQQLYLSLKAHGEGDLGTQALLIA 322

Search completed: September 16, 2004, 07:30:35
Job time : 69 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2004, 07:28:53 ; Search time 19 seconds
(without alignments)
1502.587 Million cell updates/sec

Title: US-10-067-482-2

Perfect score: 2866

Sequence: 1 MAASVRLGLVWKGKGRYP.....AKALQSDNDMSAVYRIYH 553

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Issued Patents AA:*

1: /cgn2_6/prodata/2/iaa/5A COMB.pep.*

2: /cgn2_6/prodata/2/iaa/5B COMB.pep.*

3: /cgn2_6/prodata/2/iaa/6A COMB.pep.*

4: /cgn2_6/prodata/2/iaa/6B COMB.pep.*

5: /cgn2_6/prodata/2/iaa/PCUS COMB.pep.*

6: /cgn2_6/prodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2811	98.1	547	4	US-10-164-595-72
2	733	25.6	360	4	US-09-364-230-4
3	682	23.8	345	4	US-09-364-230-8
4	537.5	18.8	247	4	US-09-364-230-2
5	364	12.7	307	4	US-09-489-039A-9294
6	343.5	12.0	299	4	US-09-328-352-7407
7	343	12.0	543	4	US-09-252-991A-27650
8	340	11.9	299	4	US-09-107-532A-5109
9	326	11.4	304	4	US-09-489-039A-9424
10	324.5	11.3	324	4	US-09-252-991A-26001
11	310	10.8	330	4	US-09-252-991A-17923
12	301.5	10.5	292	4	US-09-711-164-438
13	291	10.2	295	4	US-09-252-991A-24727
14	287.5	10.0	303	4	US-09-328-352-4879
15	266	9.3	326	4	US-09-328-352-8139
16	258.5	9.0	342	4	US-09-364-230-6
17	252	8.8	247	4	US-09-134-000C-5460
18	240.5	8.4	350	4	US-09-364-230-12
19	226.5	7.9	309	4	US-09-489-039A-9860
20	224.5	7.8	309	4	US-09-489-039A-11448
21	208	7.3	221	4	US-09-252-991A-22003
22	185.5	6.5	237	2	US-08-760-745-5
23	185.5	6.5	240	2	US-08-760-745-3
24	178.5	6.2	235	2	US-08-760-745-1
25	161.5	5.6	234	4	US-09-364-230-10
26	148.5	5.2	1360	4	US-09-788-657-22
27	145	5.1	490	4	US-09-489-039A-8350

28	139	4.8	486	4	US-09-543-681A-5380	Sequence 5380, Ap
29	138.5	4.8	316	4	US-09-252-991A-29147	Sequence 29147, A
30	135.5	4.7	786	4	US-09-509-802-2	Sequence 2, Appli
31	135.5	4.7	787	3	US-09-188-930-334	Sequence 334, App
32	135.5	4.7	787	4	US-09-312-283C-334	Sequence 334, App
33	130.5	4.6	474	3	US-09-058-692-2	Sequence 2, Appli
34	130.5	4.6	474	4	US-09-584-628-2	Sequence 2, Appli
35	127	4.4	2568	4	US-09-866-108A-3	Sequence 3, Appli
36	121.5	4.2	1024	4	US-09-562-737-49	Sequence 49, Appl
37	120	4.2	949	4	US-10-164-595-65	Sequence 65, Appl
38	120	4.2	1214	4	US-10-164-595-24	Sequence 24, Appl
39	118.5	4.1	1312	4	US-09-345-882-29	Sequence 29, Appl
40	118	4.1	474	4	US-09-134-001C-3241	Sequence 3241, Ap
41	117	4.1	723	4	US-09-252-991A-29659	Sequence 29659, A
42	113.5	4.0	289	4	US-09-107-532A-7169	Sequence 7169, Ap
43	112.5	3.9	909	3	US-09-425-383-2	Sequence 41, Appl
44	112	3.9	1005	4	US-09-206-942-41	Sequence 39, Appl
45	112	3.9	1011	4	US-09-206-942-39	Sequence 18618, A
46	111.5	3.9	720	4	US-09-252-991A-18618	Sequence 2, Appli
47	111	3.9	562	2	US-08-973-675-2	Sequence 46, Appl
48	111	3.9	605	4	US-08-714-741-46	Sequence 28, Appl
49	111	3.9	1187	1	US-08-320-559-28	Sequence 28, Appl
50	111	3.9	1187	3	US-08-545-860D-28	Sequence 28, Appl
51	111	3.9	1187	5	PCT-US94-04496-28	Sequence 28, Appl
52	111	3.9	1210	1	US-08-320-559-26	Sequence 26, Appl
53	111	3.9	1210	3	US-08-545-860D-26	Sequence 26, Appl
54	111	3.9	1210	5	PCT-US94-04496-26	Sequence 26, Appl
55	110.5	3.9	598	3	US-09-377-155-5	Sequence 5, Appli
56	110.5	3.9	598	3	US-09-377-155-13	Sequence 13, Appl
57	110.5	3.9	598	4	US-09-669-974-5	Sequence 13, Appl
58	110.5	3.9	598	4	US-09-669-974-13	Sequence 5, Appli
59	110.5	3.9	598	4	US-09-797-862-5	Sequence 5, Appli
60	110.5	3.9	598	4	US-09-797-862-13	Sequence 13, Appl
61	110.5	3.9	1581	3	US-09-110-517-2	Sequence 2, Appli
62	110	3.8	1007	4	US-09-489-039A-11327	Sequence 11327, A
63	109.5	3.8	1964	2	US-08-790-912-3	Sequence 3, Appli
64	109.5	3.8	2052	2	US-08-790-912-3	Sequence 2, Appli
65	108	3.8	909	2	US-08-363-124A-4	Sequence 4, Appli
66	108	3.8	2184	4	US-09-417-485D-6	Sequence 6, Appli
67	107.5	3.8	699	4	US-09-759-359A-2	Sequence 2, Appli
68	107	3.7	585	4	US-09-134-000C-3802	Sequence 3802, Ap
69	106.5	3.7	320	4	US-09-134-000C-5021	Sequence 5021, Ap
70	106	3.7	2860	2	US-08-826-267-2	Sequence 2, Appli
71	105.5	3.7	334	4	US-09-198-452A-923	Sequence 2, Appli
72	105.5	3.7	688	2	US-09-016-000-2	Sequence 2, Appli
73	104.5	3.6	365	2	US-08-515-251A-2	Sequence 2, Appli
74	104.5	3.6	688	3	US-09-141-047-8	Sequence 8, Appli
75	104.5	3.6	921	4	US-09-543-681A-5734	Sequence 8, Appli
76	104.5	3.6	2468	4	US-09-576-594-726	Sequence 5734, Ap
77	104	3.6	754	4	US-09-976-594-375	Sequence 375, App
78	103.5	3.6	568	1	US-08-320-559-30	Sequence 30, Appl
79	103.5	3.6	568	3	US-08-545-860D-30	Sequence 30, Appl
80	103.5	3.6	568	5	PCT-US94-04496-30	Sequence 30, Appl
81	103	3.6	204	4	US-09-198-452A-376	Sequence 376, App
82	103	3.6	277	4	US-09-489-039A-12994	Sequence 12994, A
83	102.5	3.6	594	3	US-09-377-155-7	Sequence 7, Appli
84	102.5	3.6	594	4	US-09-669-974-7	Sequence 7, Appli
85	102.5	3.6	594	4	US-09-797-862-7	Sequence 7, Appli
86	102	3.6	722	4	US-09-392-714-22	Sequence 22, Appl
87	102	3.6	912	4	US-09-540-236-1960	Sequence 1960, Ap
88	101.5	3.5	930	3	US-09-283-763-2	Sequence 2, Appli
89	101.5	3.5	930	4	US-09-574-912-2	Sequence 2, Appli
90	101.5	3.5	947	4	US-09-418-780A-1	Sequence 1, Appli
91	101.5	3.5	947	4	US-09-392-714-23	Sequence 23, Appl
92	101.5	3.5	3025	6	5223423-3	Patent No. 5223423
93	101	3.5	751	4	US-09-252-991A-27424	Sequence 27424, A
94	101	3.5	1835	3	US-08-836-325-15	Sequence 15, Appl
95	100.5	3.5	433	4	US-08-630-915A-18	Sequence 18, Appl
96	100.5	3.5	592	3	US-09-377-155-17	Sequence 17, Appl
97	100.5	3.5	592	4	US-09-669-974-17	Sequence 17, Appl
98	100.5	3.5	592	4	US-09-797-862-17	Sequence 17, Appl
99	100.5	3.5	733	3	US-08-725-459B-21	Sequence 21, Appl
100	100.5	3.5	1024	4	US-09-562-737-41	Sequence 41, Appl

ALIGNMENTS

```

RESULT 1
US-10-164-595-72
; Sequence 72, Application US/10164595
; Patent No. 6657054
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc
; TITLE OF INVENTION: Regulated Angiogenesis Genes and Polypeptides
; FILE REFERENCE: LU 103 R1
; CURRENT APPLICATION NUMBER: US/10/164,595
; CURRENT FILING DATE: 2002-06-10
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 72
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-164-595-72

Query Match      98.1%; Score 2811; DB 4; Length 547;
Best Local Similarity 98.7%; Pred. No. 3.6e-261; Indels 6; Gaps 1;
Matches 546; Conservative 0; Mismatches 1;

Qy      1  MAASLSRLGDLVWGKLGRIYPPWPGKIVNPPKDLKKPRGKCKFFVKFFGTEDHAWIKVEQL 60
Db      1  MAASLSRLGDLVWGKLGRIYPPWPGKIVNPPKDLKKPRGKCKFFVKFFGTEDHAWIKVEQL 60

Qy      61  KPYHAHKEEMIKINKGRFOQAVDAVEEFLRAKGDQTSNHSDDDKNRNSSEERSRP 120
Db      61  KPYHAHKEEMIKINKGRFOQAVDAVEEFLRAKGDQTSNHSDDDKNRNSSEERSRP 120

Qy      121  NSGDEKRLSI-SEGKVKNMGEKKRVSSESGSKGPKLRAQOSPRKGRPPKDEKD 180
Db      121  NSGDEKRLSI-SEGKVKNMGEKKRVSSESGSKGPKLRAQOSPRKGRPPKDEKD 180

Qy      181  LTIPESSTVKGMMAGPMAAFKWKQPTASEPFXDADPHFHFLLSQTEKPAVCYQAITKKLK 240
Db      181  LTIPESSTVKGMMAGPMAAFKWKQPTASEPFXDADPHFHFLLSQTEKPAVCYQAITKKLK 240

Qy      241  ICEBTGTSIQAADSTAVNGSIITPTDKKIGFLGILMGSGIVSNLLKMGHTVTVVNNRTA 300
Db      241  ICEBTGTSIQAADSTAVNGSIITPTDKKIGFLGILMGSGIVSNLLKMGHTVTVVNNRTA 300

Qy      301  EKCDLFIQEGARLGRTPAEVWSTCDITFACVSDPKAAKDLVLGPGSVLQGIKPGKCYVDM 360
Db      301  EK-----EGARLGRTPAEVWSTCDITFACVSDPKAAKDLVLGPGSVLQGIKPGKCYVDM 354

Qy      361  STVDADTVTELAAQVIVSRGGRFLEAPVSGNQQLSNDGMLVILAAAGDRGLYEDCSCFOAM 420
Db      361  STVDADTVTELAAQVIVSRGGRFLEAPVSGNQQLSNDGMLVILAAAGDRGLYEDCSCFOAM 414

Qy      421  GKTSPFLGVGNAAKMMLIVNVVQGSFMATIAEGTLAHVTQSOOTLLDIINQGLASI 480
Db      421  GKTSPFLGVGNAAKMMLIVNVVQGSFMATIAEGTLAHVTQSOOTLLDIINQGLASI 474

Qy      481  FLDDQKCONILQGNFKPDFYLYKIQDLRLAIALGDVAVNHPTPMAAAANEVYKRAKALDQS 540
Db      475  FLDDQKCONILQGNFKPDFYLYKIQDLRLAIALGDVAVNHPTPMAAAANEVYKRAKALDQS 534

Qy      541  DNDSAVRYAYIH 553
Db      535  DNDSAVRYAYIH 547

RESULT 2
US-09-364-230-4
; Sequence 4, Application US/09364230
; Patent No. 6348339
; GENERAL INFORMATION:
; APPLICANT: Caboon, Rebecca E.
; APPLICANT: Hitz, William D.
; APPLICANT: Kinney, Anthony J.
; APPLICANT: Rafalski, J. Antoni
; TITLE OF INVENTION: Enzymes Involved in Degradation of Branched-Chain Amino Acids
; FILE REFERENCE: BB-1178
; CURRENT APPLICATION NUMBER: US/09/364,230
; CURRENT FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,990
; EARLIER FILING DATE: July 31, 1998
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 8
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Glycine max
US-09-364-230-8

Query Match      23.8%; Score 682; DB 4; Length 345;
Best Local Similarity 47.1%; Pred. No. 5.7e-57;
Matches 136; Conservative 54; Mismatches 99; Indels 0; Gaps 0;

Qy      262  SITPTDKKIGFLGILMGSGIVSNLLKMGHTVTVVNNRTAEKCDLFTIQEGARLGRTPAEVV 321
Db      46  AVTEPPARIGFLGILMGSGIVSNLLKMGHTVTVVNNRTAEKCDLFTIQEGARLGRTPAEVV 105
```

```

; APPLICANT: Caboon, Rebecca E.
; APPLICANT: Hitz, William D.
; APPLICANT: Kinney, Anthony J.
; APPLICANT: Rafalski, J. Antoni
; TITLE OF INVENTION: Enzymes Involved in Degradation of Branched-Chain Amino Acids
; FILE REFERENCE: BB-1178
; CURRENT APPLICATION NUMBER: US/09/364,230
; CURRENT FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,990
; EARLIER FILING DATE: July 31, 1998
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 4
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Zea mays
US-09-364-230-4

Query Match      25.6%; Score 733; DB 4; Length 360;
Best Local Similarity 52.3%; Pred. No. 7.7e-62;
Matches 148; Conservative 47; Mismatches 88; Indels 0; Gaps 0;

Qy      266  TDKKIGFLGILMGSGIVSNLLKMGHTVTVVNNRTAEKCDLFIQEGARLGRTPAEVWSTCD 325
Db      64  SEMEVGFLGILMGKAMATNLLRHGFRVTVVNNRTLAKQCELAALGATVGETPASVSKCR 123

Qy      326  ITFACVSDPKAAKDLVLGPGSVLQGIKPGKCYVDMSTVDADTVTELAAQVIVSRGGRFLEA 385
Db      124  YTIAMLSDPSSAALSVFVKDQVLEQIGSGKGYVDMSTVDAATSKISEAVKQKGAPLEA 183

Qy      386  PVSGNQQLSNDGMLVILAAAGDRGLYEDCSCFOAMGKTSFPFLGEVGNAAKMLIVNVVQGS 445
Db      184  PVSGKXPAGDQQLVILAAAGDKPLYDGMIPAFDVLGKKSFFLGEIGNGAKMLVNVVVMG 243

Qy      446  SFMATIAEGTLAHVTQSOOTLLDIINQGLASIFLDQKCONILQGNFKPDFYLYKIQK 505
Db      244  SMNLSLEGLCLADKSGLSQTLDDVLDLJGAIANPMFKLKGPTMLQGSYSPAFPLKHQOK 303

Qy      506  DLRLAIALGDVAVNHPTPMAAAANEVYKRAKALDQSDNDMSAVY 548
Db      304  DRLALALGDENAVAMPVSAANAANFAKRSILGLGDQDQFSAVY 346

RESULT 3
US-09-364-230-8
; Sequence 8, Application US/09364230
; Patent No. 6348339
; GENERAL INFORMATION:
; APPLICANT: Caboon, Rebecca E.
; APPLICANT: Hitz, William D.
; APPLICANT: Kinney, Anthony J.
; APPLICANT: Rafalski, J. Antoni
; TITLE OF INVENTION: Enzymes Involved in Degradation of Branched-Chain Amino Acids
; FILE REFERENCE: BB-1178
; CURRENT APPLICATION NUMBER: US/09/364,230
; CURRENT FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,990
; EARLIER FILING DATE: July 31, 1998
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 8
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Glycine max
US-09-364-230-8

Query Match      23.8%; Score 682; DB 4; Length 345;
Best Local Similarity 47.1%; Pred. No. 5.7e-57;
Matches 136; Conservative 54; Mismatches 99; Indels 0; Gaps 0;

Qy      262  SITPTDKKIGFLGILMGSGIVSNLLKMGHTVTVVNNRTAEKCDLFTIQEGARLGRTPAEVV 321
Db      46  AVTEPPARIGFLGILMGSGIVSNLLKMGHTVTVVNNRTAEKCDLFTIQEGARLGRTPAEVV 105
```


GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESS: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 5109:
SEQUENCE CHARACTERISTICS:
LENGTH: 299 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...299
SEQUENCE DESCRIPTION: SEQ ID NO: 5109:
US-09-107-532A-5109
Query Match 11.9%; Score 340; DB 4; Length 299;
Best Local Similarity 28.8%; Pred. No. 3.2e-24;
Matches 76; Conservative 52; Mismatches 136; Indels 0; Gaps 0;
QY 269 KIGFLGLGMSGIVSNLLKMGHTVTVMRTAEKCDLFIQEGARLGRTPAEVYSTCDITF 328
DB 6 KLGFITGVMGSAVARHLLLEAGHEAVVNRRTAKADPLVTEGAIWADIPKVAEQSNILF 65
QY 329 ACVSDPKAAKDLVLGPGSVLQGIKPGKCYVDMSTVDADTVTBLAQVTVSRGGRFLEAPVS 388
DB 66 TWVGYKDVVEIYGGSGIFSDISGHILVLDLTSTPSLAEKIATKAKEKGADALDAPVS 125
QY 389 GNOQLSNDGMLVILAAGDRGLYEDSCSCFOAMGKTSFFLGEVGNAAKMWLVNMQGSPM 448
DB 126 GGDGLGAKNGTLTIMVGGEAVYDQVLPFKFEGTFTTTHGSAKQGHQHTKMANQIMAGTM 185
QY 449 ATTAEGTLTAHVTSQQTLLDLNQGQLASIFLDQKCNILQGNFKPDPFYLYKIQKDLR 508
DB 186 TGTEMLVYQKNGLDLKKVIETLSGSSAANWLSNYSRILKEDYTPGFFVKHFIDKDK 245
QY 509 LAIALGDVNHPTPMAAANEVYK 532
DB 246 IALEEAKMDLVLPATTAQKLYE 269

RESULT 9

US-09-252-991A-27650
Sequence 27650, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 27650
LENGTH: 543
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27650
Query Match 12.0%; Score 343; DB 4; Length 543;
Best Local Similarity 23.5%; Pred. No. 4.4e-24;
Matches 113; Conservative 83; Mismatches 212; Indels 72; Gaps 11;
QY 110 RNSSEERSRNSGDEKR-----KLSLSR---GKVKKNMGCKRVSS---GSSERGSK- 157
DB 91 RRQAGSRHPHGGDQDHPHRLPAEPHPPGDPGRQRPVPAIRHLSHADHGRRP 150
QY 158 -----SPLKRAEQSQPRKGRPP-----KDEKDLTIPESSTVYKMMGMAAFKMQPTAS 207
DB 151 GHPOKPVGGDQPRATRQPRPRTRCHRRDQLPL----- 184
QY 208 BPVKDADPH-----FHHLLSQTEKPAVCYQATYKUKLCIBETGTSIQAA 254
DB 185 -PLRAPGPHRLPGLDLRLRVQAGHH-----RSRPLDEDAQRLIRTHEQDAPASLRGAT 238
QY 255 DSTAVNGSITPDKKIGFLGLGMSGIVSNLLKMGHTVTVMRTAEKCDLFIQEGARL 313
DB 239 RQTEPGDITM---AKIGFTGTIMGKPMQAQNLQKAGHSLSLTHHDAAPADL-LEAGAIA 294
QY 314 GRTPAEVYSTCDITFACVSDPKAAKDLVLGPGSVLQGIKPGKCYVDMSTVDADTVTBLAQ 373
DB 295 LANPEVAQAEFLIWMVPTDQVEDVLPFRKDGLAEGVGPNKVVDMSSISPTATKFAE 354
QY 374 VIVSRGGRFLEAPVSGNOQLSNDGMLVILAAGDRGLYEDSCSCFOAMGKTSFFLGEVNA 433
DB 355 KIKATCAQYLDAPVSGGEGVGAATAISIMVGGCPNSFERALPLFQAMGNKNIIRVGGNGDG 414
QY 434 AKMWLVNMQGSPMATIAEGTLTAHVTSQQTLLDLNQGQLASIFLDQKCNILQGN 493
DB 415 QPAKANQIIVAINLQVAAEALLFAARNGADPAKPREALMGFPASSRILEVHGERNVKGT 474
QY 494 FKPDFYLYKIQKDLRLAIALGDVNHPTPMAAANEVYKRAKALDOSDNDMSAVRAYTH 553
DB 475 FDPGFRISLHQDLNLALAGARELNLEPNTANAQVFTSCAIGGSNDHDSALIKGLEH 534

RESULT 8
US-09-107-532A-5109
Sequence 5109, Application US/09107532A
Patent No. 6583275

US-09-489-039A-9424
 ; Sequence 9424, Application US/09489039A
 ; Patent No. 6610836
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary Breton et. al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 ; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 2709.2004001
 ; CURRENT APPLICATION NUMBER: US/09/489,039A
 ; CURRENT FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: US 60/117,747
 ; PRIOR FILING DATE: 1999-01-29
 ; NUMBER OF SEQ ID NOS: 14342
 ; SEQ ID NO 9424
 ; LENGTH: 304
 ; TYPE: PRT
 ; ORGANISM: Klebsiella pneumoniae
 US-09-489-039A-9424

Query Match 11.4%; Score 326; DB 4; Length 304;
 Best Local Similarity 29.9%; Pred. No. 7.4e-23;
 Matches 88; Conservative 52; Mismatches 136; Indels 18; Gaps 5;
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 26001
 ; LENGTH: 324
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-26001

Query Match 11.3%; Score 324.5; DB 4; Length 324;
 Best Local Similarity 26.5%; Pred. No. 1.1e-22;
 Matches 74; Conservative 62; Mismatches 142; Indels 1; Gaps 1;
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 26001
 ; LENGTH: 324
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-26001

Query Match 11.3%; Score 324.5; DB 4; Length 324;
 Best Local Similarity 26.5%; Pred. No. 1.1e-22;
 Matches 74; Conservative 62; Mismatches 142; Indels 1; Gaps 1;
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 26001
 ; LENGTH: 324
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-26001

QY 329 ACVSDPKAAKDLVLGPGSVLQGIIRPGKCYVDMSTVDADTVTTELAQVIVSRGRFLEAPVS 388
 DB 100 CCVGNDDDLRAVALGBOGAFAPGSLFVDHTTASAEVARELSLAAARELGLFLDAPVS 159
 QY 389 GNOQLSNDGMLVILAAAGDRGLYEDCSSCFQAMGKTSFFLGEVGNAAKMWLIVNMVQGSFM 448
 DB 160 GGAGAVNGALTVMVGSEAEYRAEPLLRSYARMVRKXGDSQLTKMNVQICVAGLL 219
 QY 449 ATIAEGLTLAHVTGOSQOQLDILNQGLASIFLDQKQONILQGNFKDPDFYLYKIQKDLR 508
 DB 220 QGLAEALHFARCAGLDGEAAMQVIGKGAASQWLENRHSQMLAGEFDFGFAVDNMRKDL 279
 QY 509 LAJALGDVAVNHTPMAAAANEVYKRAKALDOSDNDMSAV 547
 DB 280 ILLAARRNGAQLPVTALVDQFYAEVQAMGGGRWDTSSL 318
 RESULT 11
 US-09-252-991A-17923
 ; Sequence 17923, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 17923
 ; LENGTH: 330
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-17923

Query Match 10.8%; Score 310; DB 4; Length 330;
 Best Local Similarity 28.6%; Pred. No. 2.9e-21;
 Matches 83; Conservative 57; Mismatches 136; Indels 14; Gaps 6;
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 17923
 ; LENGTH: 330
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-17923

Query Match 10.8%; Score 310; DB 4; Length 330;
 Best Local Similarity 28.6%; Pred. No. 2.9e-21;
 Matches 83; Conservative 57; Mismatches 136; Indels 14; Gaps 6;
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 17923
 ; LENGTH: 330
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-17923

Query Match 10.8%; Score 310; DB 4; Length 330;
 Best Local Similarity 28.6%; Pred. No. 2.9e-21;
 Matches 83; Conservative 57; Mismatches 136; Indels 14; Gaps 6;
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 17923
 ; LENGTH: 330
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-17923

```
; TITLE OF INVENTION: GENES ESSENTIAL FOR MICROBIAL PROLIFERATION AND ANTISENSE THERAPY
; FILE REFERENCE: ELITRA.008A
; CURRENT APPLICATION NUMBER: US/09/711.164
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/164415
; PRIOR FILING DATE: 1999-11-9
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 438
; LENGTH: 292
; TYPE: PRN
; ORGANISM: Escherichia coli
; US-09-711-164-438

Query Match      10.2%; Score 301.5; DB 4; Length 292;
Best Local Similarity 27.7%; Pred. No. 1.6e-20;
Matches 78; Conservative 51; Mismatches 152; Indels 1; Gaps 1;

QY 269 KIGFLGLMGSGIVSNLLKMGHTVTVNRTAEKCDLFIQEGARLGRTPAEVSTCDITF 328
DB 2 KLGFIGLIMGTFMAINLARGHQHV-TTIGPVADLLSLGAVSETARQVTEASDIIF 60
QY 329 ACVSDPKAAKDLVLGPGSVLQIRPGKCYVDMSTVDADTVTELAQVIVSRGGRFLEAPVS 388
DB 61 IMVPTPQVEVLFGENGCTKASLKGKTVIVDMSSISPIETKRFARQVNNELGCDYLDAPVS 120
QY 389 GNQQLSNDGMLVILAAGDRGLYEDSCSCFOAMGKTSFFLGEVGNAAKMLIVNVQGSFMA 448
DB 121 GGEIGAREGTLIMVGGDEAVFERVKPLFELIGKNITVLVGGNGDQGTCKVANQIIIVALNI 180
QY 449 ATIAEGLTLAHVTGOSQQTLLDILNQGLASIFLDKQCNILQGNFKPDKFYLYKXIQDLR 508
DB 181 EAVSALLFASKAGADPVVRQALMGGFASSRILEVHGEMIKRTFNFGFKIALHQKDLN 240
QY 509 LAIALGDVANHPTPMAAAANEVYKRAKALDQSDNDMSAVYRA 550
DB 241 LALQAKALANLPTATCOELFNTCAANGSGQLDHSALVQA 282

RESULT 13
US-09-252-991A-24727
; Sequence 24727, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24727
; LENGTH: 295
; TYPE: PRN
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-24727

Query Match      10.2%; Score 291; DB 4; Length 295;
Best Local Similarity 28.9%; Pred. No. 1.6e-19;
Matches 82; Conservative 57; Mismatches 135; Indels 10; Gaps 5;

QY 269 KIGFLGLMGSGIVSNLLKMGHTVTVNRTAEKCDLFIQEGARLGRTPAEVST---CD 325
DB 6 KVGFIGLGGGAAMATFLVQAGLEVTVNRSAAACEPLV---ALGAARAEVGDLFGLD 61
QY 326 ITFACVSDPKAAKDLVLGPGSVLQIRPGKCYVDMSTVDADTVTELAQVIVSRGGRFLEA 385
DB 62 VVISMLADQAIRGVLL-DSGALERARPGLIHLSMSTLSLDCEALDQAHQROGLAFVAA 120
```

```
QY 386 PVSNGNQLSNDGMLVILAAGDRGLYEDSCSCFOAMGKTSFFLGEVGNAAKMLIV-NMVQ 444
DB 121 PVFGRTDVAEAGKLNIVVGGPEAIEQVKALLEIMGQKTFWFGKDPREGAVAVKISGNFMI 180
QY 445 GSWMTATIEGLTLAHVTGOSQQTLLDILNQGLASIFLDKQCNILQGNFKP-DFYLYKY 503
DB 181 ASATIESMGESVALVKRLGVBFGRFMELMSSTLFPDAPYRNYPQIVFQRTTPARFRLVIG 240
QY 504 QKDLRLAIALGDVANHPTPMAAAANEVYKRAKALDQSDNDMSAV 547
DB 241 LKQVDLALSAGKRHNVPDPLASLLHDVLLERAIHAHGDGEDSWTAL 284

RESULT 14
US-09-328-352-4879
; Sequence 4879, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4879
; LENGTH: 303
; TYPE: PRN
; ORGANISM: Acinetobacter baumannii
; US-09-328-352-4879

Query Match      10.0%; Score 287.5; DB 4; Length 303;
Best Local Similarity 25.5%; Pred. No. 3.7e-19;
Matches 74; Conservative 55; Mismatches 154; Indels 7; Gaps 2;

QY 270 IGFLGLMGSGIVSNLLKMGHTVTVNRTAEKCDLFIQEGARLGRTPAEVSTCDITFA 329
DB 10 IAFIIGLNMGRMAQNLKAGLVYGYDLSEVAIOHFAAGGIVCDSPQNAAKQADAVIT 69
QY 330 CVSDPKAAKDLVLGPGSVLQIRPGKCYVDMSTVDADTVTELAQVIVSRGGRFLEAPVS 389
DB 70 MLPAKHVKEVILGNGVLEVLKAGSLCIDSDITDPTIKDIAAQAQSKNIKICDAPVSG 129
QY 390 NQQLSNDGMLVILAAGDRGLYEDSCSCFOAMGKTSFFLGEVGNAAKMLIVNVQGSFMA 449
DB 130 GTIGAQAAGTLTPMVGCADEQTFNEVKPVLSHMGKNIVHCGDVAGAGQITAKCNLLIGISMA 189
QY 450 TIAEGLTLAHVTGOSQQTLLDILN--QGOLASIFLDKQCNILQGNFKPDKFYLYKY 502
DB 190 AVAEGVALGVKLGIDPQALAGVINTSSGRCWSDDCNPWPHINENAPASRGYQDGFATQL 249
QY 503 IQKDLRLAIALGDVANHPTPMAAAANEVYKRAKALDQSDNDMSAVYRAYI 552
DB 250 MLKDLGLAVEAAGQVKQPVILGGMVQQLYQQCMRGNALHDFSSIIQQYL 299

RESULT 15
US-09-328-352-8139
; Sequence 8139, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 8139
; LENGTH: 326
; TYPE: PRN
; ORGANISM: Acinetobacter baumannii
; US-09-328-352-8139
```

Query Match
 Best Local Similarity 9.3%; Score 266; DB 4; Length 326;
 Matches 78; Conservative 42; Mismatches 122; Indels 28; Gaps 6;

QY 270 IGFLGLMGSGIVSNLLKMGHTVTVNRTAEKCDLFTQEGA-----RLGRTPAEVVST 323
 DB 47 IAFGLGLMGSRMARLLTQAGQVAVNRTTSACEELDIGHAHLDSLNGQYP----- 100

QY 324 CDITFACVSDPKAAKDLVLGSPGVQGIIRPGKCYVDMSTVDADTVTELAQVIVSRGRFL 383
 DB 101 --LILTCLADDKAVQAVF---DQITNLKAGQVIVDFSSLSVAATKALQAASODVTVI 155

QY 384 EAPVSGNQLNDGMLVILAGDRGLYEDCSFCQAMGKTSFFLGEVGNAAKMLVNVV 443
 DB 156 DSPVSGGTGAEQGTTLVFAGGDAQTBALSPVNVLSQVTRMGDTGTGQATKICNOLI 215

QY 444 QGSPMATIAGLTLAHVTVGSOQTLIDILINQQLASIFLDKQCNIL-----QGNFKP-D 497
 DB 216 VAANSALIAEVALADRAVDVTTILAPALAG-----PADSKPQILAPRMATHTEFPVQ 270

QY 498 FYLKYIQKDLRLAIALGDAVNVHPTPMAAAA 527
 DB 271 WKVQTLKDLNNVATLANNVNDLPVAKA 300

RESULT 16
 US-09-364-230-6
 ; Sequence 6, Application US/09364230
 ; Patent No. 6348339
 ; GENERAL INFORMATION:
 ; APPLICANT: Cahoon, Rebecca E.
 ; APPLICANT: Hitz, William D.
 ; APPLICANT: Kinney, Anthony J.
 ; APPLICANT: Rafalski, J. Antoni
 ; TITLE OF INVENTION: Enzymes Involved in Degradation of Branched-Chain Amino Acids
 ; FILE REFERENCE: BB-1178
 ; CURRENT APPLICATION NUMBER: US/09/364,230
 ; CURRENT FILING DATE: 1999-07-29
 ; EARLIER APPLICATION NUMBER: 60/094,990
 ; EARLIER FILING DATE: July 31, 1998
 ; NUMBER OF SEQ ID NOS: 34
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 6
 ; LENGTH: 342
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 US-09-364-230-6

Query Match
 Best Local Similarity 9.0%; Score 258.5; DB 4; Length 342;
 Matches 81; Conservative 43; Mismatches 159; Indels 37; Gaps 6;

QY 256 STAVNGSITPTDKKIGLGLMGSGIVSNLLKMGHTVTVNRTAEKCDLFTQEGARLGR 315
 DB 25 SAAVOSQL-----ENVGFIGLGNMGHARNLNMGVYKVTVDVNVNENTMKKFSDDGIPTKL 80

QY 316 TPAEVVSTCDITFACVSDPKAAKDLVLGSPGVQGIIRPGKCYVDMSTVDADTVTE 370
 DB 81 SPLEVSXSSDVITMLPSSAHVLDVYNGRGLLAN---GGCLGPMWYIDSPVDPQTSRK 137

QY 371 LAQVI-----VSRGRFLEAPVSGNQQLSNDGMLVILAGDRGLYEDCSFCQAM 420
 DB 138 ISMDISRCTKSKKPYAEKPMMLDAPVGGVPPAEAGKTLFLVGGSEAYLAANPLLISM 197

QY 421 GXTSPFLGVBGNAKMLVNVVQGSFMATIAEGLTLAHVTVGSOQTLIDILN-----473
 DB 198 GKTYICGAGNGSVAKICNNWANGISMLGVEAFALGQNLGKASVLTDFNCSSARCW 257

QY 474 ----QGQLASIFLDKQCNILQGNFKPDPFLYKVIQKDLRLAIALGDAVNVHPTPMAAANE 529
 DB 258 SSDTYNPPGVNMDVPSSR----NYDGGFTSKMLTKDLAMASAGVGNFCPFGSQALE 313

QY 530 VYKRAKALDQSDNDMSAVYR 549
 DB 314 IYRKLCDGCELDKDFSCAFR 333

RESULT 17
 US-09-134-000C-5460
 ; Sequence 5460, Application US/09134000C
 ; Patent No. 6617156
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ; FILE REFERENCE: 032796-032
 ; CURRENT APPLICATION NUMBER: US/09/134,000C
 ; CURRENT FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: US 60/055,778
 ; PRIOR FILING DATE: 1997-08-15
 ; NUMBER OF SEQ ID NOS: 6812
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 5460
 ; LENGTH: 247
 ; TYPE: PRT
 ; ORGANISM: Enterococcus faecalis
 US-09-134-000C-5460

Query Match
 Best Local Similarity 8.8%; Score 252; DB 4; Length 247;
 Matches 58; Conservative 41; Mismatches 118; Indels 0; Gaps 0;

QY 316 TPAEVVSTCDITFACVSDPKAAKDLVLGSPGVQGIIRPGKCYVDMSTVDADTVTELAQVI 375
 DB 1 TPKAIAEASDIIITMVGSPSDVEGVYENETGIFQADLTGKIVVDLTSTPTLAEKIAKKA 60

QY 376 VSRGRFLEAPVSGNQQLSNDGMLVILAGDRGLYEDCSFCQAMGKTSFFLGEVGNAAK 435
 DB 61 AEVGAHALDAPVSGGDLGAKNGTLITMVGSDQESYDTVLPIFTGKTFMLHGSAGKGQH 120

QY 436 MMLIVNVQGSFMATIAEGLTLAHVTVGSOQTLIDILINQQLASIFLDKQCNILQGNFK 495
 DB 121 TKMANQLMIAGTMTGLTEMLVYANATGLTLEKVLTVGGSSAANWSLSNYGPRILKEDYT 180

QY 496 PDFLYKVIQKDLRLAIALGDAVNVHPTPMAAANEVYK 532
 DB 181 PGFFVNHFKDLKALDEAKKLDLPLPATQKATELYE 217

RESULT 18
 US-09-364-230-12
 ; Sequence 12, Application US/09364230
 ; Patent No. 6348339
 ; GENERAL INFORMATION:
 ; APPLICANT: Cahoon, Rebecca E.
 ; APPLICANT: Hitz, William D.
 ; APPLICANT: Kinney, Anthony J.
 ; APPLICANT: Rafalski, J. Antoni
 ; TITLE OF INVENTION: Enzymes Involved in Degradation of Branched-Chain Amino Acids
 ; FILE REFERENCE: BB-1178
 ; CURRENT APPLICATION NUMBER: US/09/364,230
 ; CURRENT FILING DATE: 1999-07-29
 ; EARLIER APPLICATION NUMBER: 60/094,990
 ; EARLIER FILING DATE: July 31, 1998
 ; NUMBER OF SEQ ID NOS: 34
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 12
 ; LENGTH: 350
 ; TYPE: PRT
 ; ORGANISM: Triticum aestivum
 ; FEATURE:
 ; NAME/KEY: UNSURE
 ; LOCATION: (154)
 ; FEATURE:
 ; NAME/KEY: UNSURE

Thu Sep 16 09:36:51 2004

Db 172 YRIGSDIGLSTVKIHHQLLAGVHIAVAEAMALAAARAGIPLTMYDVVTHAAGNSWME 231
Qy 484 QKQNTILQGNFKPDFYLYKIQKDLRLAIALGDVAVNHPTMAAAANEVYKRAKALDQSDND 543
Db 232 NRMQHVLDGDYSPKSAVDIEVKDLGLVNDTARALTPLPLATTALNMFTSASNAGFGRED 291
Qy 544 MSAYRAY 551
Db 292 DSAVIKIF 299

Query Match 8.4%; Score 240.5; DB 4; Length 350;
Best Local Similarity 26.1%; Pred. No. 1.5e-14;
Matches 80; Conservative 41; Mismatches 166; Indels 19; Gaps 6;

265 PTDKIGLGLGLMGSGIVSNLLKMGHTVTYNNRTAEKCDLFIQEGARLGRTPAEVSTC 324
Db 38 PHMESVGFGLGNMGSHMARNLVRAGYRVSVDINEVAMKFSDDGIPTKRSPLEVSESS 97
Qy 325 DITFACVSDPKAAXDLVLGPGSVL-QGIRPK-CYVDMSTVDAVTVTELAQ-----VIVS 377
Db 98 DVVITMLPSSAHVLDVYSGRLLGNGRLGFWLYIDSSTVDPTSRKISMDMGRCKLINE 157
Qy 378 RGR-----FLEAPVSGNQLSNDGMLVILAAAGDRGLYEDCSCFQAMGKTSFFLGEVGN 432
Db 158 KKGVAEKPIMDAPVPGGXPKAKVGHXXXXXXXKXXXXXXXXXXLLANGKKLIYCGGAGN 217
Qy 433 AAKMMLIVNVQGSFMATIAEGLTLAHVTGQSQQTLDDLNOGQLASIFLD--QKQNIL 490
Db 218 GSAKLCNNMAWASMLGVSEAFALGQNLGKASTLTDIENCSSARCWSSDTYNPVGVM 277
Qy 491 QG-----NFKPDFYLYKIQKDLRLAIALGDVAVNHPTMAAAANEVYKRAKALDQSDNDMS 545
Db 278 TGVPSRNVGGFTSKLMAKDLDLAMASAGVGYKCPMGSEALEYRKLCDGECFFKDFS 337
Qy 546 AVYRAY 551
Db 338 CAPRHF 343

RESULT 19
US-09-489-039A-9860
; Sequence 9860, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9860
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9860

Query Match 7.9%; Score 226.5; DB 4; Length 309;
Best Local Similarity 22.4%; Pred. No. 2.7e-13;
Matches 69; Conservative 64; Mismatches 160; Indels 15; Gaps 6;

Qy 249 TSIQAADSTAVNGSITPTDKKIGLGLMGSGIVSNLLKMGHTVTYVW--NRTAEKCDLF 306
Db 2 TNVHKETSMHAHTNVC-----VIGLSMGMAARACLAQ--LNTWGVDPNDNCRAL 52
Qy 307 IOEGARLGRTPAEV--VSTCDITFACVSDPKAAXDLVLGPGSVLQIRPKCYVDMSTVD 364
Db 53 LAAGAN-GAGPSAVPFAAEALDAVLLVNAQAQVGRILFGESGLAAHLKPTGVVMSSTIA 111
Qy 365 ADVTTELAQVIVSRGGRFLEAPVSGNQLSNDGMLVILAAAGDRGLYEDCSCFQAMGKTS 424
Db 112 SADAQAIAEALAEYQVLLMLDPAFVSGGAVKAAAGMTVMASGDAAFARLAPVLDAVAGV 171
Qy 425 FFLG-EVGNAAKMLIVNVQGSFMATIAEGLTLAHVTGQSQQTLDDLNOGQLASIFLD 483
Db 172 YRIGSDIGLSTVKIHHQLLAGVHIAVAEAMALAAARAGIPLTMYDVVTHAAGNSWME 231
Qy 484 QKQNTILQGNFKPDFYLYKIQKDLRLAIALGDVAVNHPTMAAAANEVYKRAKALDQSDND 543
Db 232 NRMQHVLDGDYSPKSAVDIEVKDLGLVNDTARALTPLPLATTALNMFTSASNAGFGRED 291
Qy 544 MSAYRAY 551
Db 292 DSAVIKIF 299

RESULT 21
US-09-252-991A-22003
; Sequence 22003, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18

```

, PRIOR APPLICATION NUMBER: US 60/094,190
, PRIOR FILING DATE: 1998-07-27
, NUMBER OF SEQ ID NOS: 33142
, SEQ ID NO 22003
, LENGTH: 221
, TYPE: prt
, ORGANISM: pseudomonas aeruginosa
US-09-252-991A-22003

```

Query Match	7.3%;	Score 208;	DB 4;	Length 221;
Best Local Similarity	27.5%;	Pred. No. 9.5e-12;		
Matches 53;	Conservative 36;	Mismatches 104;	Indels 0;	Gaps 0;
Qy	270	IGFGLGLMGSGIYVNLKMGHTVTYVNNRTAEKCDLFQIEGARLGRTPAEVVSVCIDITFA	329	
Db	9	IATFGLGNMGGPMANLLKAGHRVNFVLDQPKAVLSVEQAGADSALQCCGAEVVIS	68	
Qy	330	CVSDPKAAKDVLGPGSVLGQIRPGKCYVDMSTVDADTVTLAGIVSRGRGFLEAPVSG	389	
Db	69	MLPAGQHVESLYLGDGGLARVAGKPLLLDCSTTAPETARKVAEAAAKGLTLLDAPVSG	128	
Qy	390	NQQLSNDGMLVILAAGDRGLYEDCSSCFQAMGKT'SFFLPGVGNAAKMWLVNMGVQGSFMA	449	
Db	129	GVGGARAGTTSFVCGGPAEGFARARPVLENNGRNIFHAGDHGAGQVAKICNNMLLGIIMA	188	
Qy	450	TIAPGLTLAHVTG	462	
Db	189	GTAEALALGVKNG	201	

RESULT 22
US-08-760-745-5
Sequence 5, Application US/08760745
Patent No. 5972658
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Goli, Lyrna K.
APPLICANT: Murry, Lynn E.
TITLE OF INVENTION: LUNG GROWTH FACTOR VARIANT
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08760,745
FILING DATE: Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PP-0169 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 237 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

```
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 945419
US-08--760-745-5

Query Match          6.5%; Score 185.5; DB 2; Length 237;
Best Local Similarity 29.8%; Pred.No.1.5e-09;
Matches 65; Conservative 32; Mismatches 94; Indels 27; Gaps 9;

QY      7 RLGDLVWGKLGRYPPWPGKIWNPPKDLKKPRGKKCFVFKEGTEDHAWIKVEQLKPYHAH 66
       ||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db      11 KCGDLVFAKWKGYPHWFARIDEMPAVKSTANK-YQVFFFGTHETAF LGPKDLPFYEES 69
       ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||

QY      67 KEEMIKINKGRFQOAVDAVEEFLLRAKGDQTSSHNSSDDKNRRNSSEERSRPSNGDEK 126
       ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db      70 KEKGFKENKRKGFSGLWEIEN-----NPTVKASGYOSSQKSCAAFEPEVEPEAHEGGDG 124
       ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||

QY      127 RKLSLSEGKVKKNMGECKRKRVSSSGSSRGSKSPDKR-----AQOSPRKRGPPPK----D 177
       ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db      125 KKGS-AGGSDE---EGKLVIDEPAKENKGT LKRRAGDVLEDSPKR----PKESGDHE 176
       ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||

QY      178 EKDLTIPESSTVKGMAGPMAAFKWQPTASEPVKDADP 215
       ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db      177 EED---KEIAALEGERHLPEVEK-NSTPSEPDSGGQG 210
       ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||

RESULT 23
US-08-760-745-3
; Sequence 3, Application US/08760745
; Patent No. 5972658
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Golii, Suriya K.
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: LUNG GROWTH FACTOR VARIANT
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/760,745
; FILING DATE: Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0169 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 240 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; LIBRARY: GenBank
; CLONE: 598956
US-08-760-745-3
```


Db 74 SI-----EEFVDSLEKPRILLVMKAGEA-TDKTIAALTPLHDKGILIDGG 119
Qy 380 GRFLEAPVSGNQQLSNDGMLV-----ILAAQDGLYEDCSCCFQAMGK 422
Db 120 NTFFKDTIRNRRELSAQGFNFIGTVSGGEGALKGPSIMPGQKEAYELVAPILEKIAA 179
Qy 423 TS-----FFLGEVGNAAKMLIVNMVQGSFMATIAEGLT-LAHVTGQSQTLLDIL--- 472
Db 180 VADGEPCTYIGADGAGHYVQVHNGIEYQMDLIAEAYSVLKHSGLJTNELADTFTEM 239
Qy 473 NOGQIASIFLQKONIL-----QGNFKDPFLYKIQKDLRLAIALGDAVNHPTPMAAAA 527
Db 240 NKGELSS-YLIEITADIFRKKDEGNLYVD-----VILDEANKGT----- 279
Qy 528 NEVYKRAKALQSDNDM-----SAVYRAYI 552
Db 280 -----GKWTQSLSLDLGPVTLITESVFARYI 306

RESULT 29

US-09-252-991A-29147
; Sequence 29147, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29147
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29147

Query Match 4.8%; Score 138.5; DB 4; Length 316;
Best Local Similarity 24.3%; Pred. No. 8.1e-05;
Matches 44; Conservative 28; Mismatches 102; Indels 7; Gaps 2;
Qy 378 RGRFLEAPVSGNQQLSNDGMLVILAAQDGLYEDCSCCFQAMGKTSFELGEVGNAAKMM 437
Db 130 RGLAMDAPVSGGTAGAAAGTLTFMVGGDAEALAKARPLFEAWGRNIFHAGDPDGAGQVAK 189
Qy 438 LIVNMVQGSFMATIAEGLT-LAHVTGQSQTLLDILNQGLASIFLD--QKQNILQG--- 492
Db 190 VCNQQLAVLMTGTAEMALGVANGLEAKVLAEIMRSGGNWALEVYNPFWGVENAPA 249
Qy 493 --NFKPDPFLYKIQKDLRLAIALGDAVNHPTPMAAAANEVYKRAKALQSDNDMGAIVYRA 550
Db 250 SRDYGSGFMAQLMAKDLGLAQBAAQASASTPMGSLALSRLYLLKQGYAERDFSVQKL 309
Qy 551 Y 551
Db 310 F 310

RESULT 30

US-09-509-802-2
; Sequence 2, Application US/09509802
; Patent No. 6489130
; GENERAL INFORMATION:
; APPLICANT: Immunex Corp.
; APPLICANT: Bird, Timothy
; APPLICANT: Virga, G.D.
; TITLE OF INVENTION: DEATH ASSOCIATED KINASE CONTAINING ANKYRIN REPEATS (DAKAR)
; FILE REFERENCE: 2889-US
; CURRENT APPLICATION NUMBER: US/09/509,802

; CURRENT FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 786
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-509-802-2

Query Match 4.7%; Score 135.5; DB 4; Length 786;
Best Local Similarity 20.0%; Pred. No. 0.0007;
Matches 126; Conservative 83; Mismatches 233; Indels 189; Gaps 30;
Qy 11 LVWGLKGRYPWPWG-----KIVN-----PPKOLKKPRGKKCFVFKFGTGDHAWI 55
Db 212 VIWGLVITQKPFADKKNILHIMKVKVKGHRPELPP--ICRPRPRAC--ASLIGLMQRCWH 267
Qy 56 KVEQLKPYHAHKEMIKNGKRFQOAVDAVEFELARRAKGDKQDTSSHNSDDKNRPNSE 115
Db 268 ADPQVRP-----TFQBITSETDLCEKPDDEYKDLAHEFGKSSLESKE 312
Qy 116 ERSRPNSGDEKR-----KLSLSEGVKVKNNMGEGK 144
Db 313 --ARPESSRLKASAPPDNDCSLSSELLSQDLSISQTLGPEELSRSSSECKLPSSSG 370
Qy 145 KRVS-----SGSSERGSKPLKRAEQSPKRGPRPPKDEKDLTIPESSTVKGMMAGPMA 198
Db 371 KRLSGVSSVDSAFSSRGSLS--LSPEREASTGDLG--PTD-----IQKKLVDAIISGDT 422
Qy 199 AFK--WQPTASEPVKADPHFHHLLSQTEKPAVCYQAITKKLKICEETGTSIQAAADS 256
Db 423 RLMKILQPDVDLVLDDSSASLLHLAVEAGQECVKWLLNNANPNLTNRKGSTPLHMA-- 480
Qy 257 TAVNGSIPTDKKIGFLGLGMLGSGIVSNLLKMGHTV-----TVNRTAEKCD--- 304
Db 481 -----VERK-----GRGIVELLARKTSVNAKDEQDWTALHFAQNGDEAST 522
Qy 305 -LFIQEGARL-----GRTPAEVWSTCDITTFACVSDPKAAKDLV--LQPSGVYLGIRPK 355
Db 523 RLLLEKNASVNEVDPEGTPMHV-----AC---QHQQENIVRTLRRRGVDVGLQGD 571
Qy 356 CYVDMSTVDADTVTELAQVIVSRGGRFLEAPVSGNQQLSNDGMLVILAAQDGLY----- 410
Db 572 AMLPLHYAAWQGHLPVIVLLAKQPG-----VSVNAQ--TLDGRTPLHLAAQGRHYRVARI 624
Qy 411 -----EDCSCFOAMGKTSFFLGEVGNAAKMMLI VNMVQGSFMATIAEGLT-LAHVTGQSQ 465
Db 625 LIDLCSDVNICSLOQOTPLHVAAETGHTSTARLLHARGAK--EALTSEGYTALHAAQN- 682
Qy 466 QTLDLILNQGLASI--FLDQKQONILQGNFKDPFLYKIQKDLRLAIALGDAVNHPTPM 523
Db 683 -----GHLATVKLLIEEKADVMARGPLN-----QTALHAAARGH----- 717
Qy 524 AAAANEVYKR---AKALQSDND--MSAVYRA 550
Db 718 ----SEVVEELVSDALIDLSDPEQGLSALHLA 744

Search completed: September 16, 2004, 07:29:23
Job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2004, 07:29:27 ; Search time 718 Seconds
(without alignments)

247.335 Million cell updates/sec

Title: US-10-067-482-2

Perfect score: 2866

Sequence:

1 MAAVSLRLGLVWGKLGKLYP.....AKALDQSDNMSAVRAYIH 553

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1342398 seqs, 321133274 residues

Total number of hits satisfying chosen parameters: 13423398

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2866	100.0	553	US-10-067-482-2	Sequence 2, Appli
2	2811	98.1	547	US-10-067-482-3	Sequence 3, Appli
3	2798	97.6	550	US-10-103-313-417	Sequence 417, App
4	1410	49.2	276	US-10-067-482-4	Sequence 4, Appli
5	1404	49.0	269	US-10-103-313-540	Sequence 540, App
6	1291	45.0	249	US-09-987-755-2	Sequence 2, Appli
7	1291	45.0	260	US-10-103-313-474	Sequence 474, App
8	1286	44.9	250	US-10-103-313-317	Sequence 317, App
9	740.5	25.8	364	US-10-425-114-50561	Sequence 50561, A
10	727	25.4	293	US-10-425-114-50561	Sequence 174476
11	720	25.1	290	US-10-167-547C-14	Sequence 14, Appl
12	695	24.2	289	US-10-167-547C-16	Sequence 16, Appl
13	682	23.8	333	US-10-424-599-144794	Sequence 144794
14	569	19.9	343	US-10-437-963-116017	Sequence 116017
15	567	19.8	256	US-10-425-114-42840	Sequence 42840, A

16	494.5	17.3	199	US-10-437-963-174477	Sequence 174477, A
17	375	13.1	292	US-10-282-122A-52083	Sequence 52083, A
18	367	12.8	299	US-09-912-020-256	Sequence 256, App
19	354.5	12.4	344	US-10-767-701-42795	Sequence 42795, A
20	350	12.2	296	US-10-282-122A-56894	Sequence 56894, A
21	348.5	12.2	296	US-10-156-761-9081	Sequence 9081, Ap
22	347	12.1	344	US-10-437-963-151834	Sequence 151834, A
23	340	11.9	295	US-10-282-122A-57891	Sequence 57891, A
24	334.5	11.7	295	US-10-123-965B-11	Sequence 11, Appl
25	331	11.5	286	US-10-282-122A-60763	Sequence 60763, A
26	331	11.5	305	US-10-156-761-9563	Sequence 9563, Ap
27	330	11.5	285	US-10-282-122A-59047	Sequence 59047, A
28	330	11.5	285	US-10-335-977-5655	Sequence 5655, Ap
29	327.5	11.4	290	US-10-282-122A-44990	Sequence 44990, A
30	324.5	11.3	288	US-10-282-122A-66301	Sequence 66301, A
31	318	11.1	296	US-10-282-122A-66289	Sequence 66289, A
32	314.5	11.0	277	US-10-282-122A-50872	Sequence 50872, A
33	314	11.0	297	US-10-282-122A-67475	Sequence 67475, A
34	313.5	10.9	301	US-10-282-122A-52906	Sequence 52906, A
35	305	10.6	350	US-10-424-599-255872	Sequence 255872, A
36	301.5	10.5	292	US-10-282-122A-42674	Sequence 42674, A
37	301.5	10.5	292	US-10-287-274-438	Sequence 438, App
38	297.5	10.4	356	US-10-425-114-53506	Sequence 53506, A
39	296	10.3	176	US-10-767-701-60840	Sequence 60840, A
40	291.5	10.2	292	US-10-282-122A-74927	Sequence 74927, A
41	290.5	10.1	298	US-10-282-122A-61452	Sequence 61452, A
42	287.5	10.0	215	US-10-424-599-146727	Sequence 146727, A
43	282	9.8	318	US-10-282-122A-73480	Sequence 73480, A
44	281.5	9.8	291	US-10-282-122A-73480	Sequence 73480, A
45	281	9.8	292	US-10-043-487-224	Sequence 224, App
46	281	9.8	336	US-10-282-122A-75520	Sequence 75520, A
47	279.5	9.8	292	US-10-408-765A-755	Sequence 755, App
48	279.5	9.8	305	US-10-408-765A-755	Sequence 116817, A
49	278	9.7	284	US-10-437-963-116817	Sequence 437, A
50	277	9.7	302	US-10-238-075-301	Sequence 301, App
51	272.5	9.5	292	US-10-282-122A-45255	Sequence 45255, A
52	272	9.5	288	US-10-282-122A-45255	Sequence 45255, A
53	269	9.4	297	US-10-282-122A-47679	Sequence 47679, A
54	256	8.9	121	US-10-767-701-60846	Sequence 60846, A
55	255	8.9	255	US-10-335-977-5654	Sequence 5654, Ap
56	254.5	8.9	552	US-10-437-963-114005	Sequence 114005, A
57	229.5	8.0	1354	US-10-437-963-163633	Sequence 163633, A
58	206.5	7.2	154	US-10-767-701-61555	Sequence 61555, A
59	205	7.2	81	US-10-767-701-52389	Sequence 52389, A
60	199.5	7.0	305	US-10-282-122A-51175	Sequence 51175, A
61	190.5	6.6	202	US-09-866-050A-651	Sequence 651, App
62	189	6.6	203	US-09-768-826-41	Sequence 41, Appl
63	189	6.6	203	US-09-833-245-2264	Sequence 2264, Ap
64	189	6.6	203	US-10-247-671-132	Sequence 132, App
65	187.5	6.5	274	US-10-425-114-69397	Sequence 69397, A
66	186.5	6.5	246	US-09-925-302-511	Sequence 511, App
67	186.5	6.5	246	US-09-925-302-511	Sequence 511, App
68	185.5	6.5	221	US-10-418-445-4	Sequence 4, Appli
69	185.5	6.5	237	US-09-938-885A-5	Sequence 5, Appli
70	185.5	6.5	240	US-09-938-885A-3	Sequence 3, Appli
71	185.5	6.5	240	US-09-987-755-7	Sequence 7, Appli
72	185.5	6.5	240	US-10-257-021-14	Sequence 14, Appl
73	185.5	6.5	240	US-10-207-791-2	Sequence 2, Appli
74	185.5	6.5	240	US-10-116-275-220	Sequence 220, App
75	185	6.5	240	US-10-424-599-243466	Sequence 243466
76	181.5	6.3	305	US-10-156-761-8219	Sequence 8219, Ap
77	178.5	6.2	235	US-09-938-885A-1	Sequence 1, Appli
78	176	6.1	356	US-10-424-599-164051	Sequence 164051, A
79	170.5	5.9	326	US-10-276-774-2486	Sequence 2486, Ap
80	169	5.9	314	US-09-768-826-43	Sequence 43, Appl
81	169	5.9	314	US-09-833-245-2266	Sequence 2266, Ap
82	169	5.9	670	US-09-823-187-86	Sequence 86, Appl
83	169	5.9	670	US-09-863-776-51	Sequence 51, Appl
84	169	5.9	671	US-09-946-374-308	Sequence 308, App
85	169	5.9	671	US-09-823-187-85	Sequence 85, Appl
86	169	5.9	671	US-09-863-776-20	Sequence 20, Appl
87	169	5.9	671	US-10-206-915-346	Sequence 346, App
88	169	5.9	671	US-10-199-670-346	Sequence 346, App

89 169 5.9 671 12 US-10-201-858-346 Sequence 346, App
90 169 5.9 671 12 US-10-205-890-346 Sequence 346, App
91 169 5.9 671 12 US-10-208-024-346 Sequence 346, App
92 169 5.9 671 12 US-10-201-853-346 Sequence 346, App
93 169 5.9 671 12 US-10-174-581-346 Sequence 346, App
94 169 5.9 671 12 US-10-176-483-346 Sequence 346, App
95 169 5.9 671 12 US-10-176-749-346 Sequence 346, App
96 169 5.9 671 12 US-10-176-914-346 Sequence 346, App
97 169 5.9 671 12 US-10-176-915-346 Sequence 346, App
98 169 5.9 671 12 US-10-006-485A-308 Sequence 308, App
99 169 5.9 671 12 US-10-013-907A-308 Sequence 308, App
100 169 5.9 671 12 US-10-015-499A-308 Sequence 308, App

ALIGNMENTS

RESULT 1
US-10-067-482-2
; Sequence 2, Application US/10067482
; Publication No. US20030148407A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc.
; TITLE OF INVENTION: Angiogenesis Dehydrogenase Gene
; FILE REFERENCE: 1U 102 R1
; CURRENT APPLICATION NUMBER: US/10/067,482
; CURRENT FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 553
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-067-482-2

Query Match 100.0%; Score 2866; DB 14; Length 553;
Best Local Similarity 100.0%; Pred. No. 3e-235;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAAVSLRLGDLVWGKLGRIYPPWPGKI VNPDKLKKPRGKKCFVFFGTEHAWIKVEQL 60
DB 1 MAAVSLRLGDLVWGKLGRIYPPWPGKI VNPDKLKKPRGKKCFVFFGTEHAWIKVEQL 60
QY 61 KPYHAHKEEMIKINKGRFQQAVDAVEEFLLRAKGDQTS SHNSDDKNRNSSEERSRP 120
DB 61 KPYHAHKEEMIKINKGRFQQAVDAVEEFLLRAKGDQTS SHNSDDKNRNSSEERSRP 120
QY 121 NSGDEKRLSLSEGKVKKNMGEGKKRVSSGSSERGSKSPKRAEQSPKRGPRPKDEKD 180
DB 121 NSGDEKRLSLSEGKVKKNMGEGKKRVSSGSSERGSKSPKRAEQSPKRGPRPKDEKD 180
QY 181 LTIPESSVTKGMAGPMAAFKMQPTASEPVKDADPHFHHFLLSQTEKPAVCYQAITKKLK 240
DB 181 LTIPESSVTKGMAGPMAAFKMQPTASEPVKDADPHFHHFLLSQTEKPAVCYQAITKKLK 240
QY 241 ICEEETGTSIOAADSTAVNGSIPTDCKIGFLGLMGSGIVSNLLKMGHTVTVWNRTA 300
DB 241 ICEEETGTSIOAADSTAVNGSIPTDCKIGFLGLMGSGIVSNLLKMGHTVTVWNRTA 300
QY 301 EKCDLFIQEGARLGRTPAEVSTCDITFACVSDPKAAKDLVLGPGSVLGQIRPGKCYVDM 360
DB 301 EKCDLFIQEGARLGRTPAEVSTCDITFACVSDPKAAKDLVLGPGSVLGQIRPGKCYVDM 360
QY 361 STVDADTVTELAQVIVSRGRFLEAPVSGNQQLSNDGMLVILAAGDRGLYEDCSSCFQAM 420
DB 361 STVDADTVTELAQVIVSRGRFLEAPVSGNQQLSNDGMLVILAAGDRGLYEDCSSCFQAM 420
QY 421 GKTSFPLGEGVNAAKMMLIVNNVQGSFMATIAEGLTLAHTVGOSQOTLLDILNQGLASI 480
DB 421 GKTSFPLGEGVNAAKMMLIVNNVQGSFMATIAEGLTLAHTVGOSQOTLLDILNQGLASI 480
QY 481 FLDDQCKNTLQGNFKPDPFYLYKI QKDLRLAIALGDVANNHPTMAAANEVYKRAKALDOS 540

DB 481 FLDDQCKNTLQGNFKPDPFYLYKI QKDLRLAIALGDVANNHPTMAAANEVYKRAKALDOS 540
QY 541 DNDMSAVYRAYIH 553
DB 541 DNDMSAVYRAYIH 553
RESULT 2
US-10-067-482-3
; Sequence 3, Application US/10067482
; Publication No. US20030148407A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc.
; TITLE OF INVENTION: Angiogenesis Dehydrogenase Gene
; FILE REFERENCE: 1U 102 R1
; CURRENT APPLICATION NUMBER: US/10/067,482
; CURRENT FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 547
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-067-482-3

Query Match 98.1%; Score 2811; DB 14; Length 547;
Best Local Similarity 98.7%; Pred. No. 1.4e-230;
Matches 546; Conservative 0; Mismatches 1; Indels 6; Gaps 1;
QY 1 MAAVSLRLGDLVWGKLGRIYPPWPGKI VNPDKLKKPRGKKCFVFFGTEHAWIKVEQL 60
DB 1 MAAVSLRLGDLVWGKLGRIYPPWPGKI VNPDKLKKPRGKKCFVFFGTEHAWIKVEQL 60
QY 61 KPYHAHKEEMIKINKGRFQQAVDAVEEFLLRAKGDQTS SHNSDDKNRNSSEERSRP 120
DB 61 KPYHAHKEEMIKINKGRFQQAVDAVEEFLLRAKGDQTS SHNSDDKNRNSSEERSRP 120
QY 121 NSGDEKRLSLSEGKVKKNMGEGKKRVSSGSSERGSKSPKRAEQSPKRGPRPKDEKD 180
DB 121 NSGDEKRLSLSEGKVKKNMGEGKKRVSSGSSERGSKSPKRAEQSPKRGPRPKDEKD 180
QY 181 LTIPESSVTKGMAGPMAAFKMQPTASEPVKDADPHFHHFLLSQTEKPAVCYQAITKKLK 240
DB 181 LTIPESSVTKGMAGPMAAFKMQPTASEPVKDADPHFHHFLLSQTEKPAVCYQAITKKLK 240
QY 241 ICEEETGTSIOAADSTAVNGSIPTDCKIGFLGLMGSGIVSNLLKMGHTVTVWNRTA 300
DB 241 ICEEETGTSIOAADSTAVNGSIPTDCKIGFLGLMGSGIVSNLLKMGHTVTVWNRTA 300
QY 301 EKCDLFIQEGARLGRTPAEVSTCDITFACVSDPKAAKDLVLGPGSVLGQIRPGKCYVDM 360
DB 301 EKCDLFIQEGARLGRTPAEVSTCDITFACVSDPKAAKDLVLGPGSVLGQIRPGKCYVDM 360
QY 361 STVDADTVTELAQVIVSRGRFLEAPVSGNQQLSNDGMLVILAAGDRGLYEDCSSCFQAM 420
DB 361 STVDADTVTELAQVIVSRGRFLEAPVSGNQQLSNDGMLVILAAGDRGLYEDCSSCFQAM 420
QY 421 GKTSFPLGEGVNAAKMMLIVNNVQGSFMATIAEGLTLAHTVGOSQOTLLDILNQGLASI 480
DB 421 GKTSFPLGEGVNAAKMMLIVNNVQGSFMATIAEGLTLAHTVGOSQOTLLDILNQGLASI 480
QY 481 FLDDQCKNTLQGNFKPDPFYLYKI QKDLRLAIALGDVANNHPTMAAANEVYKRAKALDOS 540
DB 481 FLDDQCKNTLQGNFKPDPFYLYKI QKDLRLAIALGDVANNHPTMAAANEVYKRAKALDOS 540
QY 541 DNDMSAVYRAYIH 553
DB 541 DNDMSAVYRAYIH 547
RESULT 3
US-10-103-313-417
; Sequence 417, Application US/10103313

```
; Publication No. US20030082758A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P207C1
; CURRENT APPLICATION NUMBER: US/10/103,313
; CURRENT FILING DATE: 2002-03-12
; PRIORITY APPLICATION: 2002-03-12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 417
; LENGTH: 550
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-103-313-417

Query Match          97.6%; Score 2798; DB 14; Length 550;
Best Local Similarity 98.7%; Pred. No. 1.9e-229;
Matches 543; Conservative 0; Mismatches 1; Indels 6; Gaps 1;

Qy 4 VSLRLGDLVWGKLGRIYPPWPKIVNPPKDLKKPRGKCFVKKPFGTGDHAWIKVEQLKPY 63
Db 7 VSLRLGDLVWGKLGRIYPPWPKIVNPPKDLKKPRGKCFVKKPFGTGDHAWIKVEQLKPY 66

Qy 64 HAHKEEMIKINKGRFQQAQVDAVEEFLLRAKGDQTSNNSDDKNRNSSEERSRNSG 123
Db 67 HAHKEEMIKINKGRFQQAQVDAVEEFLLRAKGDQTSNNSDDKNRNSSEERSRNSG 126

Qy 124 DEKKLSLSGKVKKNNGEKKRVSSSGSSRGSKSPKRAQEQSPKRGPPPKDEKDLTI 183
Db 127 DEKKLSLSGKVKKNNGEKKRVSSSGSSRGSKSPKRAQEQSPKRGPPPKDEKDLTI 186

Qy 184 PESTVTKGMAGPMAAPKWOPTASEPVKQADPHPHFLLSQTEKPAVCYQAITKKLICE 243
Db 187 PESTVTKGMAGPMAAPKWOPTASEPVKQADPHPHFLLSQTEKPAVCYQAITKKLICE 246

Qy 244 EETGTSIQAADSTAVNGSITPTDKTGFLGLGMLGSGGIVSNLLKMGHTVTVNRTAEK 303
Db 247 EETGTSIQAADSTAVNGSITPTDKTGFLGLGMLGSGGIVSNLLKMGHTVTVNRTAEK 305

Qy 304 DLFIQEGARLGRTPAEVVSTCDITFACVSDPKAAKDLVLGSPGVLOGIRGKCYVDMSTV 363
Db 306 -----EGARLGRTPAEVVSTCDITFACVSDPKAAKDLVLGSPGVLOGIRGKCYVDMSTV 360

Qy 364 DADVTVELAQVIVSRGRGFLEAPVSGNQQLSNDGMLVILAAAGDRGLYEDCSSCFQAMGKT 423
Db 361 DADVTVELAQVIVSRGRGFLEAPVSGNQQLSNDGMLVILAAAGDRGLYEDCSSCFQAMGKT 420

Qy 424 SFPLGEVGNAAKMWLI VNVVQSGFMATIAEGLTLAHVTGSGSQOTLLDILNQGLASIFLD 483
Db 421 SFPLGEVGNAAKMWLI VNVVQSGFMATIAEGLTLAQVTGSGSQOTLLDILNQGLASIFLD 480

Qy 484 OKQNILQGNFKDPDFYLYKIQDLRLAIALGDVNHPTPMAAANAEVYKRAKALDQSDND 543
Db 481 OKQNILQGNFKDPDFYLYKIQDLRLAIALGDVNHPTPMAAANAEVYKRAKALDQSDND 540

Qy 544 MSAVYRAYIH 553
Db 541 MSAVYRAYIH 550

RESULT 4
US-10-067-482-4
; Sequence 4, Application US/10067482
; Publication No. US20030148407A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc.
; TITLE OF INVENTION: Angiogenesis Dehydrogenase Gene
; FILE REFERENCE: IU 102 R1
; CURRENT APPLICATION NUMBER: US/10/067,482
; CURRENT FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 4
; LENGTH: 276
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-067-482-4

Query Match          49.2%; Score 1410; DB 14; Length 276;
Best Local Similarity 100.0%; Pred. No. 1.4e-111;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 278 MGSIVSNLLKMGHTVTVNRTAEKCDLFIQEGARLGRTPAEVVSTCDITFACVSDPKAA 337
Db 1 MGSIVSNLLKMGHTVTVNRTAEKCDLFIQEGARLGRTPAEVVSTCDITFACVSDPKAA 60

Qy 338 KDLVLGSPGVLOGIRGKCYVDMSTVDADVTVELAQVIVSRGRGFLEAPVSGNQQLSNDG 397
Db 61 KDLVLGSPGVLOGIRGKCYVDMSTVDADVTVELAQVIVSRGRGFLEAPVSGNQQLSNDG 120

Qy 398 MLVILAAAGDRGLYEDCSSCFQAMGKTSFFLGEVGNAAKMWLI VNVVQSGFMATIAEGLTL 457
Db 121 MLVILAAAGDRGLYEDCSSCFQAMGKTSFFLGEVGNAAKMWLI VNVVQSGFMATIAEGLTL 180

Qy 458 AHVTGSGSQOTLLDILNQGLASIFLDQKNILQGNFKDPDFYLYKIQDLRLAIALGDV 517
Db 181 AHVTGSGSQOTLLDILNQGLASIFLDQKNILQGNFKDPDFYLYKIQDLRLAIALGDV 240

Qy 518 NHPTPMAAANAEVYKRAKALDQSDNDMSAVYRAYIH 553
Db 241 NHPTPMAAANAEVYKRAKALDQSDNDMSAVYRAYIH 276

RESULT 5
US-10-103-313-540
; Sequence 540, Application US/10103313
; Publication No. US20030082758A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P207C1
; CURRENT APPLICATION NUMBER: US/10/103,313
; CURRENT FILING DATE: 2002-03-12
; NUMBER OF SEQ ID NOS: 653
; PRIORITY APPLICATION: 2002-03-12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 540
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: (4)_feature
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (165)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-103-313-540

Query Match          49.0%; Score 1404; DB 14; Length 269;
Best Local Similarity 99.3%; Pred. No. 4.5e-111;
Matches 267; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 VSLRLGDLVWGKLGRIYPPWPKIVNPPKDLKKPRGKCFVKKPFGTGDHAWIKVEQLKPY 63
Db 1 VSLRLGDLVWGKLGRIYPPWPKIVNPPKDLKKPRGKCFVKKPFGTGDHAWIKVEQLKPY 60

Qy 64 HAHKEEMIKINKGRFQQAQVDAVEEFLLRAKGDQTSNNSDDKNRNSSEERSRNSG 123
Db 61 HAHKEEMIKINKGRFQQAQVDAVEEFLLRAKGDQTSNNSDDKNRNSSEERSRNSG 120

Qy 124 DEKKLSLSGKVKKNNGEKKRVSSSGSSRGSKSPKRAQEQSPKRGPPPKDEKDLTI 183
Db 121 DEKKLSLSGKVKKNNGEKKRVSSSGSSRGSKSPKRAQEQSPKRGPPPKDEKDLTI 180
```



```

; Sequence 50561, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 50561
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3136-019-H11_FLI pep
US-10-425-114-50561

```

Query Match	25.8%;	Score 740.5;	DB 12;	Length 364;
Best Local Similarity	47.8%;	Pred. No. 2.6e-54;		
Matches 160;	Conservative 54;	Mismatches 104;	Indels 17;	Gaps 5

Qy	216	FHFHFLLSQTKEP- AVCYQAITKK-LKICEEETGTSIQAADSTAVNGSIITPTDKKIGFL	273
Dd	31	KHK-----TRTGP SAYTYRGETRSAVRACEE-----GTBEEG-CQOSEMEVGEL	75
Qy	274	GLGLMGSGIVNLKUMGH TVTNRTARKCDLFIQEGARLGRTPAEVSTCDDITFACVSD	333
Dd	76	GLGIMGXAMATNLLRHG FRVTWNRNTLAKCQELAAALGATVGETPASVSVSKCRYTIAMLSD	135
Qy	334	PKAAKDLVLGSPGYLQGR PKCKVCYDMSDTVADTVTELAQVIVSRGGRFLEAPVSGNOQL	393
Dd	136	PSAALSVPFDKGVLQEIG SGKGYDMSTVDAATSTKISEAVKQKGGAFLEAPVSGSKXP	195
Qy	394	SNDGMVLILAAGDRGLYE DCSSCFQAMGKTSSFGEVGNAAKMMLIANNVQGSFWATTAE	453
Dd	196	AENGQLVILAAGDPLDGM TPADFVLGKGFLEIGEINGCARMKLVNNVMGSMNNSLSE	255
Qy	454	GLTLAHVTGOSQQTLDB LDLNQGLASIFLDQCKQNILQCNFKPDPLYKIOKDRLALAL	513
Dd	256	GLCLADKSGLSQPTLLDV LDLGAIANPMFKLGFTMLQGSISPAFFLKHXQKDXMELALAL	315
Qy	514	GDAVNHTPMAAAAENEVY KRAKALDQSDNDMSAVY	548
Dd	316	GDENAVAMPYSAAANEAF KKARSJGLGDODPSAVY	350

RESULT 10
US-10-437-963-174476
; Sequence 174476, Application US/10437963
; Publication No. US20040123343A1

```

: GENERAL INFORMATION:
: APPLICANT: La Rosa, Thomas J.
: APPLICANT: Kovalic, David K.
: APPLICANT: Zhou, Yihua
: APPLICANT: Cao, Yongwei
: APPLICANT: Wu, Wei
: APPLICANT: Boukharov, Andrey A.
: APPLICANT: Barbazuk, Brad
: APPLICANT: Li, Ping
: TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
: TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
: FILE REFERENCE: 38-21(53221)B
: CURRENT APPLICATION NUMBER: US/10/437,963
: CURRENT FILING DATE: 2003-05-14
: NUMBER OF SEQ ID NOS: 204966
: SEQ ID NO 174476
: LENGTH: 293

```

```

; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_72412C.1.pcp
US-10-437-963-174476

```

Query Match 25.4%; Score 727; DB 16; Length 293;
Best Local Similarity 52.1%; Pred. No. 2.7e-53;
Matches 146; Conservative 49; Mismatches 85; Indels 0; Gaps 0;

[illegible]

RESULT 11

```

US-10-167-547C-14
; Sequence 14, Application US/10167547C
; Publication NO. US20030170653A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont de Nemours and Company
; APPLICANT: Damude, Howard G.
; TITLE OF INVENTION: A Biological Method for the Production
; TITLE OF INVENTION: Butyrolactone and its Intermediates
; FILE REFERENCE: CL1804 US NA
; CURRENT APPLICATION NUMBER: US/10/167,547C
; CURRENT FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: 60/297198
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Microsoft Office 07
; SEQ ID NO 14
; LENGTH: 290
; TYPE: PRT
; ORGANISM: tulip pistil
US-10-167-547C-14

```

Query Match 25.1%; Score 720; DB 14; Length 290;
Best Local Similarity 50.7%; Pred. No. 1e-52;
Matches 142; Conservative 52; Mismatches 86; Indels 0; Gaps 0;

QY	269	KIGFLGLGMSGIVSNLLKNGHTVTVNNRTAEKCDLFIQEGARLGRTPAEVVSTCDITF	328
Db	2	EVGFLGLGIMKAMAVNLLRSGFRVTNNRTLSKCNELLEQGASVGETPAAVTKKCKYTT	61
QY	329	ACVSDPKAAKDLVLGPGSVLQGRPCYCYDMSTVDADTVTELAQIVTSRSGRFLPAPYS	388
Db	62	AWLSDPSNALSVPFDKGVLEHMSGKGYIDMSTVDAVTSSKISEAITQKGGHFLEAPYS	121
QY	389	GNQSLNDGMLVILAAGDRGLYEDSCSCFQAMGKTGFLLGEVNAAKMLIVNMQGSFM	448
Db	122	GSKPAEDGQVILAAGEKALYEITPAFEVLGKSGFFLGQVNGFANMKLIVNMINGSM	181
QY	449	ATIAEGLTAAHTVQSQOQLDILNQGLASIFLDQCKNIIQGNFKPDPYLKYIQKDLR	508
Db	182	NALSEGUSLAGSGUEQKTLDDVLIDGATANPMFKLGPANTQNNHPAPFLKHQCKDMR	241

Qy	482	L D Q K C O N I L Q G F P D F Y L K Y T Q K L R L A I A L G D V A N H P T M A A A A N E V Y K R A K A L D Q S D	541
D _b	267	F S L K G P S M V K A Y P T A F L K H O O K L R L A L A E S V S Q S I P T V A A N E L Y K V A K S I G L A D	326
Qy	542	N D M S A V Y R A	550
D _b	327	O D F S A V I E A	335

```

RESULT 15
US-10-425-114-42840
; Sequence 42840, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kowalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-211531313B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 42840
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700580946_F1.1.bep
US-10-425-114-42840

```

Query Match	19.8%;	Score 567;	DB 12;	Length 256;
Best Local Similarity	46.1%;	Pred. No. 9.6e-40;		
Matches 111;	Conservative 51;	Mismatches 79;	Indels 0;	Gaps 0;
Qy	310	GARLGRTPAEVYSTCDITFEACVSDPKAAKDLVLGSPVLOGIRPGKCVDVMSTVDADT	369	
Db	5	GAKYEPTPAQVASSCDVTFAMLADPQSAAEVACSSGAAGLAPGKGVVDVSTVDGATSK	64	
Qy	370	ELAQVIVSRGRFLFAPVSGNQQLNSDGMVLVLAAGDRGLYEDSCSCFQAWGKTSFFLGE	429	
Db	65	LIGRITSTGASFLFLEAVSVSGKPAEDGLLIFLTAGDESLYKRVAPLLDVMGSRFLGD	124	
Qy	430	VGNAAKMMLVNMVQGSFMATIAEGLTLAHVTGOSQDTLLDILNQGLASIFLDQKQCN	489	
Db	125	VGNGAAMKLVNMVGMVMVVSFSEGLLLSEKVLDPNTLVEVISOAISAPMSELKQPSM	184	
Qy	490	LOGNFKPDFVLKYLOKDLRLAIALGDVAHNPTPMAAAANEVYKAKALDOSDNDMSAVYR	549	
Db	185	VKAAYPPAFPLKQXQOKDLRLALAAESVQSOSIPTVAAANELYKAAKSLGLSDHDFS	244	
Qy	550	A 550		
Db	245	A 245		

RESULT 16
US-10-437-963-174477
; Sequence 174477, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping

```

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221) B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 174477
; LENGTH: 199
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(199)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_72413C.1.pep
; US-10-437-963-174477

```

Query Match	17.3%	Score	494.5	DB	16	Length	199
Best Local Similarity	52.3%	Pred. No.	9.9e-34				
Matches	101	Conservative	31	Mismatches	60	Indels	1
Gaps	1						
QY	287	LKGHTVTVMNRTAEKCDLFQEGARLCRTPAEYVSTCDITFACVSDPKAAKDVLVGPSS	346				
DB	1	LEHGRVTVMNRTLSKCOELVALGNAVGETPAAVAKRYTIAMLSDFSALSVVFDKDG	60				
QY	347	VLOGIRPGCKYVDMSTVDADTWTETLAQVIVSRGGRFLFAPVSGNQOGLNDGMLVILAAGD	406				
DB	61	VLEQITGEGKGVDMSTVDAATSCISEAIKQGGAFVEAPVSGSKKPAEDGQLVILAAGD	120				
QY	407	RGLYEDSCSCFOAMGKTSFFFLGVEGVNRAKMWLVNMVQSGSMATTAEGLTAHVTVGQSQO	466				
DB	121	KVLVDDMVPAFDVLGKKSFFLGEIGNGAHMKLVNMIXG-MNNALUSEGLSADNLSGLSPQ	179				
QY	467	TLLDILNQQLAS	479				
DB	180	TLFDVLDTGAIAN	192				

RESULT 17

US-10-282-122A-52083

; Sequence 52083, Application US/10282122A

; Publication No. US20040029129A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari

; APPLICANT: Zyskind, Judith

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John

; APPLICANT: Carr, Grant

; APPLICANT: Yamamoto, Robert

; APPLICANT: Forsyth, R.

; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

RESULT 20
US-10-282-122A-56894
; Sequence 56894, Application US/10282122A

```

; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CI7849_1.pep
US-10-767-701-42795

Query Match      12.4%; Score 354.5; DB 16; Length 344;
Best Local Similarity 29.7%; Pred. No. 2e-21;
Matches 86; Conservative 54; Mismatches 145; Indels 5; Gaps 3

QY      246 TGSTSIQAAADSTAVNGS---ITPTDKIGIFGLGLMGSGIVSNLLKMGHTVTVMNRATK 302
DB      23 TAAAAAAMSSSSSTAANVSDEPISPDITTRVAWVGTVGMVQSMAGHLLSAGYALTAVFNRTASK 82

QY      303 CDLFIQEGARLGRTPAEVVSCTDITFACVSDPKAAKDLVLGPS-VLQGIIRPKCVVDMS 361
DB      83 TQSLVSRGASLADSPRAAASAADVTFLMVGFPSDVRSTALDPESTGALSLAGPGIIVDMT 142

QY      362 TVDADTVTELAQVIVSRGGRFLEAFVSGNQQLSNDGMLVILLAGDRGLEYEDSCSFCQAMG 421
DB      143 TSDPTLAEIAIAAAAAAGCSA VDAFVSGDGRGARNAAALSIFAGGDAAVVARLAPLFLKMG 202

QY      422 KTSFPLGEVGNAAKMWLVNMVQGSFMATIAEGLTLAHVTGOSQQTLLDLINQOGLASIF 481
DB      203 -NALYMGFGAGQRAKLGNGTAIASTMTVLGVBMGVYAHKAGLDVAKWLEBRISTIGAAGSKS 261

QY      482 LDQKQNTILQGNFKPDFYLKTYIKDLRLAIALGDVANNHPTPMAAAANEVY 531
DB      262 LELYKRIILERDMAAGFYVRHVPKDLGICLSECQAMGLSLPGLAQAQVLY 311

RESULT 20
US-10-282-122A-56894
; Sequence 56894 , Application US/10282122A

```



```
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_51941C.1.pap
US-10-437-963-151834

Query Match      12.1%; Score 347; DB 16; Length 344;
Best Local Similarity 27.5%; Pred. No. 8.6e-21;
Matches 84; Conservative 65; Mismatches 149; Indels 8; Gaps 4;

QY 233 QAITKK---LKICEETGTSIQAADSTAVNGS----IPTDKKIGFLGLMGSGIVSNL 286
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 7 RSLTRRRPPLPSAAAAAAMSSATGNNVSDRPISPDITTRVAVWGTVGVMQSMAGHL 66
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 287 LKMGHTVWNRATAEKCDLFIQEGARLGRTPAEVSTCDITFACVSDPKAAKDLVGPS- 345
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 67 LAAGYALIVNRTASKAEGLSRGATLAESPRAAAAAADVIFLMVGFSDVRSTSLDPST 126
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 346 GVLQIRPGKCVDMSTVDADTVTTELAQVIVSRGRFLEAPVSGNQQLSNDGMLVILAAG 405
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 127 GALAGLAFGLLVDMTTSPTLAAEIAEAAAASCAAVDAPVSGDGRGARSATLSIFAGG 186
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 406 DRGLYEDCSSCFQAMGKTSFFFLGVEGNAAKMMLIVNMVQGSFMAIABGLTLAHVTGASQ 465
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 187 DAAVVARLAPLPLKMG-NALYMGFGAGQRAKLGNGQIAIATSMVGLVEGMVYAHKAGLDV 245
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 466 QTLIDLINQQLASIFLDQKQCNILQGNFKPDVFLKYIOKDLRLALALGDVAVNHPTEMAA 525
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 246 AKWLSAISTGAAGSKSLDLYGKRMLEKRDMAAGFYVRHFVKDLGICLISECQAMGLALPGLA 305
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 526 AANEVY 531
      : : : : :
Db 306 LAHQLY 311
      : : : : :
```

```
RESULT 23
US-10-282-122A-57891
; Sequence 57891, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Lianguo
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
```

```
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 57891
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Enterococcus faecium
US-10-282-122A-57891

Query Match      11.9%; Score 340; DB 12; Length 295;
Best Local Similarity 28.8%; Pred. No. 2.7e-20;
Matches 76; Conservative 52; Mismatches 136; Indels 0; Gaps 0;

QY 269 KIGFLGLMGSGIVSNLLKMGHTVTVNRTAEKCDLFIQEGARLGRTPAEVSTCDITF 328
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2 KLGFIGTGVMSGSAVARHLEAGHEAVNRTAKADPLVTEGAIWADTPKAVAEOSNLLF 61
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 329 ACVSDPKAAKDLVGLPGSVLQGIREFKCYVDMSTVDADTVTTELAQVIVSRGRFLEAPVS 388
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 62 TMVGYPKDVEEIIYVQSGIFSADISGHILVDLTSTPSLAEKIAKTAKEKGADALDAPVS 121
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 389 GNOQLSNDGMLVILAAGDRGLYEDCSSCFQAMGKTSFFFLGVEGNAAKMMLIVNMVQGSFM 448
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 122 GGDLAGKNGTTLIMVGGBEAVVDVLPFKFEGTFTTLHGSAGKGQHTKVAQNIIMAGTM 181
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 449 ATIAEGLTLAHVTGOSQOTLLDLINQQLASIFLDQKQCNILQGNFKPDVFLKYIOKDLR 508
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 182 TGTEMLVVAQKNGLDLKKVIETLSGGSAAWNSLSNYSRILLKEDYTFGFFVVKHFKDKL 241
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 509 LAIALGDVAVNHPTEMAAANEVYK 532
      : : : : :
Db 242 IALEEAEKMDLVLPATTQALKLYE 265
      : : : : :
```

```
RESULT 24
US-10-123-965B-11
; Sequence 11, Application US/10123965B
; Publication No. US20030099966A1
; GENERAL INFORMATION:
; APPLICANT: Huang, Sheng-He
; TITLE OF INVENTION: E. Coli Virulence Determinants and Methods of Use Thereof
; FILE REFERENCE: USPI841A-SHH
; CURRENT APPLICATION NUMBER: US/10/123,965B
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/284,762
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 11
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-123-965B-11
```

```
Query Match      11.7%; Score 334.5; DB 14; Length 295;
Best Local Similarity 28.7%; Pred. No. 7.9e-20;
Matches 81; Conservative 53; Mismatches 147; Indels 1; Gaps 1;

QY 270 IGFLGLMGSGIVSNLLKMGHTVTVNRTAEKCDLFIQEGARLGRTPAEVSTCDITFA 329
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3 IGFITGIMGKPMAYNLOQAGHTLYFSAHFEPAPQEFIGERGIVCSTPTPEVAQECVIT 62
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 330 CVSDPKAAKDLVGPS-GVLOGIRPGKCVDMSTVDADTVTTELAQVIVSRGRFLEAPVS 388
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 63 MLPDTPHVEDVLFHPNYGVHLSHGKIVIDMSSISPVATKAPAQRIIIVAGAEVIDAPVS 122
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 389 GNOQLSNDGMLVILAAGDRGLYEDCSSCFQAMGKTSFFFLGVEGNAAKMMLIVNMVQGSFM 448
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 123 GBEVGAKAGTSLSIMVGGCEVYLQIKPILELMGKNITLVGNVGGDTCKVANOIIVALNI 182
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 449 ATIAEGLTLAHVTGOSQOTLLDLINQQLASIFLDQKQCNILQGNFKPDVFLKYIOKDLR 508
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

APPLICANT: Ollsen, Nall
APPLICANT: Zyskind, Judith

RESULT 28
US-10-335-977-5655
; Sequence 5655, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS

Query Match	11.5%;	Score 330;	DB 12;	Length 285;
Best Local Similarity	32.7%;	Pred. No. 1.Be-19;		
Matches	92;	Conservative 46;	Mismatches 137;	Indels 6; Gaps 4;
Qy	269	KIGFLGILGMSGIVSNLLKMGHTVTWNNRTAEKCDLFIQEGARLGRTPAEWSVSTCDITF	328	
Db	2	KIGWIGLGANGTPMATFLRDLAGLEVSVYNRTESKAAPLKEKGAVVYTS:PIDLAAKVDLVF	61	
Qy	329	ACVSDPKAAKXDLVLGPGSGVLQGRTPGKCYVDMSTVDADTVELAQVTVSRGGRLEAPVS	388	
Db	62	TMLSD-KAAAYDAVLAPKFEWQMSK-KIVNMNMSTIAFLPESLSLEKIAQKHQATYLEAPVS	118	
Qy	389	GNQQLSNDGMVLILAAAGDRGLYEDCSFCFOAMGKTSFPLGEVGNAAKMLIVNNVQGSFM	448	
Db	119	GSVCAAKAGALLILAAAGDEEVI SKLKEPVLAHLSQTFYLVKGVGQGTGAKLSINSLAQMG	178	
Qy	449	ATTABGLTLAHVTGOSQOITLLDIINQOLASIFLDQKCNILQGNFPKDFYLVKVIQKDLR	508	
Db	179	VAYSEALLAKRLGVDAESFLQIIGSGMNSPLFQAKKGWMLQNYPAAPSLKLMUKDIR	238	
Qy	509	LA-IALGDVAHNPHTPMAAAAANEVYKRAKALDQSDNDMSAVY	548	
Db	219	LAKNEAGEAMK-LPFLFOAEELVYSOAEKSGILGGLDMAVY	277	

RESULT 29

US-10-282-1228-44990

; Sequence 44990, Application US/10282122A

; Publication No. US20040029129A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2004, 07:28:53 ; Search time 19 seconds
(without alignments)
2799.680 Million cell updates/sec

Title: US-10-067-482-2

Perfect score: 2866

Sequence: 1 MAAVSLRLGLVWGKLGKLYP.....AKALDQSDNDMSAVIRYIH 553

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

PIR 78:*
1: piri:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	392.5	13.7	288	2	C70303
2	392	13.7	289	2	G90314
3	384	13.4	334	2	T08967
4	375	13.1	292	2	G97310
5	368	12.8	288	2	B69870
6	367	12.8	299	1	JQ0613
7	367	12.8	299	2	C91129
8	367	12.8	299	2	C85974
9	361	12.6	294	2	AE0897
10	338	11.8	290	2	S76381
11	335	11.7	291	2	AB3380
12	334	11.7	296	2	T34859
13	333	11.6	286	2	AC1558
14	331.5	11.6	299	2	B83979
15	331	11.5	286	2	AE1200
16	330	11.5	285	2	G71912
17	329	11.5	289	2	B81065
18	329	11.5	289	2	F81402
19	324.5	11.3	288	2	E83447
20	318.5	11.1	315	2	H82512
21	318	11.1	296	2	H83456
22	318	11.1	315	2	A98338
23	318	11.1	315	2	AI2944
24	316	11.0	293	2	G86901
25	316	11.0	371	2	T10635
26	310	10.8	296	2	C83232
27	309	10.8	294	2	B96015
28	303.5	10.6	301	2	G96013
29	301.5	10.5	292	2	D64782
30	301.5	10.5	297	2	E96736
31	299.5	10.5	292	2	B90700
32	299.5	10.5	292	2	E85550
33	295	10.3	289	2	E90248
34	295	10.3	299	2	D96736
35	293	10.2	291	2	D83371
36	293	10.2	293	2	G95277
37	291.5	10.2	298	2	AI0947
38	290	10.1	346	2	A32867
39	288	10.0	286	2	E87444
40	287.5	10.0	298	2	S40826
41	281.5	9.8	298	2	E91229
42	281.5	9.8	298	2	D86076
43	279.5	9.8	292	2	AB0567
44	279.5	9.8	292	2	JC7926
45	279	9.7	298	2	C96022
46	278	9.7	313	2	AH3058
47	278	9.7	313	2	H98227
48	270	9.4	300	2	AH2912
49	270	9.4	300	2	C97687
50	269.5	9.4	302	2	AI3645
51	264.5	9.2	298	2	B83553
52	263.5	9.2	310	2	AB3338
53	261.5	9.1	298	1	C42902
54	259.5	9.1	292	2	AG2225
55	259.5	9.1	299	2	T18682
56	254.5	8.9	295	2	A96019
57	253	8.8	304	2	AC3064
58	253	8.8	304	2	H98222
59	252	8.8	295	2	H95295
60	245.5	8.6	296	2	AI0443
61	244.5	8.5	1197	2	D86317
62	239	8.3	302	2	D65054
63	234.5	8.2	294	2	B70825
64	224.5	7.8	261	2	A69807
65	213.5	7.4	295	2	C87417
66	209.5	7.3	298	2	E70707
67	203	7.1	287	2	E72666
68	201.5	7.0	310	2	H75510
69	198	6.9	301	2	B64164
70	189	6.6	203	2	JC7163
71	186.5	6.5	237	2	JC5660
72	185.5	6.5	240	2	A55055
73	168.5	5.9	1392	2	T51947
74	162	5.7	472	2	AD1609
75	162	5.7	472	2	AH1246
76	160.5	5.6	669	2	JC5662
77	159.5	5.6	530	2	JC7168
78	157.5	5.5	304	2	S59414
79	154.5	5.4	476	2	AC2465
80	149	5.2	1445	2	T50508
81	148.5	5.2	1360	2	JC5839
82	146.5	5.1	632	2	T48316
83	146	5.1	468	2	AI4250
84	141.5	4.9	484	2	T01658
85	141	4.9	468	2	D56146
86	140	4.9	468	2	162463
87	139.5	4.9	511	2	T05363
88	139	4.8	468	2	I84555
89	139	4.8	468	2	AI4249
90	139	4.8	468	2	F90982
91	139	4.8	468	2	D85828
92	138.5	4.8	492	2	T25520
93	137	4.8	1095	2	T25520
94	136.5	4.8	468	1	DEEGC
95	136	4.7	468	1	DEEGC
96	135.5	4.7	483	2	D8167
97	135.5	4.7	486	2	T44750
98	135.5	4.7	2364	2	A56577
99	135	4.7	469	2	A72377
100	134.5	4.7	508	2	T42523

probable dehydroge
carbonic semialde
probable oxidoredu
3-hydroxyisobutyra
probable dehydroge
probable dehydroge
Probable D-threoni
probable oxidoredu
3-hydroxyisobutyra
conserved hypothet
hypothetical 31.2K
probable dehydroge
probable dehydroge
2-hydroxy-3-oxopro
3-hydroxyisobutyra
probable 3-hydroxy
2-HYDROXY-3-OXOPRO
3-hydroxyisobutyra
3-hydroxyisobutyra
oxidoreductase [imp
D-threonine dehydr
3-hydroxybutyrate
probable 3-hydroxy
3-hydroxyisobutyra
3-hydroxyisobutyra
3-hydroxyisobutyra
3-hydroxyacid dehy
3-hydroxyisobutyra
probable dehydroge
3-hydroxyisobutyra
probable 3-hydroxy
probable 3-hydroxy
protein F15H18.21
hypothetical prote
probable 3-hydroxy
3-hydroxyisobutyra
3-hydroxyisobutyra
3-hydroxyisobutyra
hypothetical prote
probable 3-hydroxy
3-hydroxyisobutyra
hypothetical prote
hepatoma-derived g
hepatoma-derived g
hepatoma-derived g
probable transcrip
6-phosphogluconate
6-phosphogluconate
hepatoma-derived g
lens epithelium-de
hypothetical prote
6-phosphogluconate
hypothetical prote
GTP-N protein - h
hypothetical prote
phosphogluconate d
phosphogluconate d
phosphogluconate d
phosphogluconate d
phosphogluconate d
phosphogluconate d
gluconate-6-phosph
gluconate-6-phosph
6-phosphogluconate
hypothetical prote
phosphogluconate d
phosphogluconate d
6-phosphogluconate
probable phosphogl
microtubule-associ
6-phosphogluconate
probable phosphogl

ALIGNMENTS

RESULT 1
C70303
3-hydroxyisobutyrate dehydrogenase - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 11-Jun-1999
C:Accession: C70303
V:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Oviatt, D.; et al. 1998. The complete genome of the hyperthermophilic bacterium Aquifex aeolicus. Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666; PMID:9537320
A:Accession: C70303
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-288 <AOF>
A:Cross-references: GB:AE000670; NID:g2982779; PIDN:AA06408.1; PID:g2982783; GB:AE00065
A:Experimental source: strain VF5
C:Genetics:
A:Gene: htd
A:Superfamily: 3-hydroxyisobutyrate dehydrogenase; 3-hydroxyisobutyrate dehydrogenase htd
P:3-265/Domain: 3-hydroxyisobutyrate dehydrogenase homology <HIB>
Query Match 13.7%; Score 392.5; DB 2; Length 288;
Best Local Similarity 31.7%; Pred. No. 9e-19;
Matches 90; Conservative 61; Mismatches 132; Indels 1; Gaps 1;
QY 269 KIGFLGLGMSGIVSNLLKMGHTVTVNRTAEKCDLFIQ-
DB 2 KVGFIGLHGLRAIAKRLIEGQVELIVNRTLSKAHEFAKTGAEVTSPPADLINKVDV 61
QY 328 FACVSDPKAAKDLVLGSGVLQGIKPGKCYVDMSTVDADTVTELAQVIVSRGRFLEAPV 387
DB 62 FVIVFDSQAESEVIFGEKGLVKGDIKGTVIDMTNHYLYAKAVEELKCLGAFYLDAPV 121
QY 388 SGNQOLSNDGMLVILAAAGDRLGYEDCCSFQAMGKTSFFLGEVGNAAKMLIVNVQGSF 447
DB 122 LGSVIPALKGELTIVGGDKFENKPLFKFCAIYLLGAGMSKMLIVNVLGGI 181
QY 448 MATTAEGTLTAHVTSQSOQTLLDILNOGLASIFLDQKQNILOQNFNFPDFVLYKIQKDL 507
DB 182 MEVLAETALGEKAGIDKETIIVNLNDGAGKSYLDLVKKLLEDFSTHESVNLVYKDL 241
QY 508 RLATLALGDVAVNHPTPMAAAANEVYKRAKALDQSDNDMSAVYRAY 551
DB 242 HYAQLIKDLGLFSFTTAAVKETGLARKEGFGNLDPSAVYKLF 285
RESULT 2
G90314
oxidoreductase [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 15-Jun-2001
A:Accession: G90314
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: G90314
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-289 <KUR>
A:Cross-references: GB:AE006641; NID:g13814790; PIDN:AAK41774.1; GSPDB:GN00155
C:Genetics:
A:Gene: SSO1560
C:Superfamily: 3-hydroxyisobutyrate dehydrogenase; 3-hydroxyisobutyrate dehydrogenase htd
Query Match 13.7%; Score 392; DB 2; Length 289;
Best Local Similarity 29.3%; Pred. No. 9.8e-19;

Matches 83; Conservative 68; Mismatches 132; Indels 0; Gaps 0;
QY 269 KIGFLGLGMSGIVSNLLKMGHTVTVNRTAEKCDLFIQ-
DB 2 KVGFIGLHGLRAIAKRLIEGQVELIVNRTLSKAHEFAKTGAEVTSPPADLINKVDV 61
QY 329 ACVSDPKAAKDLVLGSGVLQGIKPGKCYVDMSTVDADTVTELAQVIVSRGRFLEAPV 388
DB 62 SMWTDAPVEVLEFVGVNGVSKNRLGILFVDMSTNSPEFAKVTIKRLSEYGMFLDAPV 121
QY 389 GNOOLSNDGMLVILAAAGDRLGYEDCCSFQAMGKTSFFLGEVGNAAKMLIVNVQGSF 448
DB 122 GSKGAREGLTITMVGKEDVFKEVPEPIFKAMGNIIHVGDVSGQALKLCNQVVVALNM 181
QY 449 ATTAEGTLTAHVTSQSOQTLLDILNOGLASIFLDQKQNILOQNFNFPDFVLYKIQKDL 508
DB 182 VSVVEGELLARSIGIDDDKLFVSLTGAANSFTVQYLYPKIMKGDNLGPFKAAHLKKDLK 241
QY 509 LALATLGDVAVNHPTPMAAAANEVYKRAKALDQSDNDMSAVYRAY 551
DB 242 YAMEIANSKSLPLGTSIALQLYNAMVSLGIGELGTQGLVKVY 284
RESULT 3
T08967
hypothetical protein F19B15.150 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 18-Aug-2000
A:Accession: T08967
R:Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft, C.; et al. 1999. The Arabidopsis genome. Nature 392, 253-256, 1999
A:Reference number: 216519
A:Accession: T08967
A:Molecule type: DNA
A:Residues: 1-334 <BEV>
A:Cross-references: EMBL:AL078470; GSPDB:GN00062; ATSP:F19B15.150
A:Experimental source: cultivar Columbia; BAC clone F19B15
C:Genetics:
A:Gene: ATSP:F19B15.150
A:Map position: 4
A:Superfamily: 3-hydroxyisobutyrate dehydrogenase; 3-hydroxyisobutyrate dehydrogenase htd
F:40-301/Domain: 3-hydroxyisobutyrate dehydrogenase homology <HIB>
Query Match 13.4%; Score 384; DB 2; Length 334;
Best Local Similarity 32.2%; Pred. No. 4e-18;
Matches 96; Conservative 48; Mismatches 152; Indels 2; Gaps 2;
QY 254 ADSTAVNGSIPTDKKIGFLGLGMSGIVSNLLKMGHTVTVNRTAEKCDLFIQ-
DB 24 ASSTISSDIITPSNTKIGWIGVMSGCHLIRAGYVTVFNRTISKAQTLIDMGANV 83
QY 314 GRTPAEVVTCDITFACVSDPKAAKDLVLG-
DB 84 ADSPNSVAEQSDVFTIVGYPSDVHVLDDPKSGALSGRLRGVVLVDMTSPSLAE 143
QY 373 QVIVSRGRFLEAPVSGNQQLSNDGMLVILAAAGDRLGYEDCCSFQAMGKTSFFLGEVGN 432
DB 144 KAASFNCFSIDAPVSGGDLGAKNKLISFAGGDETTVKRLDPLFSLMGKNV-FMGTS 202
QY 433 AAKMLIVNVQGSFPMATIEGLTIAHVTSQSOQTLLDILNOGLASIFLDQKQNILOQ 492
DB 203 GQFALANQITITASTMLGLVEGLIYAHKAGLDVKKFLEAISTGAAGSKSIDLYGRILKR 262
QY 493 NEFKPDFYLYKIQKDLRLAALGDVAVNHPTPMAAAANEVYKRAKALDQSDNDMSAVYRA 550
DB 263 FDFPDGFVNHVFKDLGICLINECQRMGLALPGLALAAQLYLSLKAHEGEGDLGTQALL 320

RESULT 4
G97310
dehydrogenase related to 3-hydroxyisobutyrate dehydrogenase, YKWC B. subtilis ortholog
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001

C:Accession: G97310
R:Nolling, J.; Brennon, G.; Omelchenko, M.V.; Markatova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium ClC
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: G97310
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-292 <KUR>
A:Cross-references: GB:AEO01437; PIDN:AAK81274.1; PID:g15026424; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC3342
C:Superfamily: 3-hydroxyisobutyrate dehydrogenase; 3-hydroxyisobutyrate dehydrogenase hc

Query Match 13.1%; Score 375; DB 2; Length 292;
Best Local Similarity 30.0%; Pred. No. 1.3e-17;
Matches 85; Conservative 55; Mismatches 143; Indels 0; Gaps 0;

QY 269 KIGFLGLGMSGGVISNLKMGHTVTWNRKTAEKCDLFIQEGARLGRTPAEWSCTDITP 328
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 7 KIGFIGTGMVGKMIKNLLKGGYVHVYARTKEALPLVEIGAVLETISKELSORCVII 66
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

QY 329 ACVSDPKRAKDVLPGSGVLOGIRPKCYVDSTVDADTVTELAQVIVSRGRLEAPVS 388
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 67 TMVGYPDSVDEEYFNDNGILANLKPOSYVVDMTTPSLAKKIYAACKERNIFALDAPVS 126
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

QY 389 GNQQLSNDGMLVILAAGRGLYEDCSSCFQAMGKTSEFFGEVNAAKMVLINVMQGSM 448
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 127 GGDIGAABGTLSIMVGGDKAEFDALMPIELMKGNTVLQGSAGSGGHTRKCNQIAIASNI 186
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

QY 449 ATTAEGLTLAHVTGOSQOTLLDLMOGLASTFLDQCKNILQGNFKPDFLVKYIKDLR 508
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 187 MGVCALSYAKKSGLNASTVLTKSIGAGGAASQWLSAYAFRLAGDFNPGEYIHKFVKDMK 246
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

QY 509 LAIALGDVAHNPTMAAANEVYKRAKALDQSDNDMSAVYRAY 551
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 247 IALEEADEMGLKTPALELSKKLYDLKXGKEDLGTOALYHLY 289
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

RESULT 5
B69870
3-hydroxyisobutyrate dehydrogenase homolog ykwc - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
C:Accession: B69870
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Bouilliet, S.; Bruschii, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chc
A.: Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Fulgider, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
diech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y., M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, M.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terptrata, P.; Togouchi, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yanahane, K.; Yasunoto, K.; Yata, K.; Yoshida, K.
A:Authors: Yoshikawa, H.F.; Zemanstein, E.; Yoshikawa, H.; Darchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: B69870
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-288 <KUN>
A:Cross-references: GB:P299111; GB:AL009126; MID:g2633699; PIDN:CAB13269.1; PID:g2633767
A:Experimental source: strain 168
C:Genetics:
A:Gene: ykwc
C:Superfamily: 3-hydroxyisobutyrate dehydrogenase; 3-hydroxyisobutyrate dehydrogenase hc
F.5-266/Domain: 3-hydroxyisobutyrate dehydrogenase homology <HB>

Qy 509 LAIALGDAVNHTPMMAAANEVYKRAKALDQSDNDMSAV 547
Db 242 NALDTSHGVGAQLPLTAIVMMEMQALRADGHNDHSHAL 280

RESULT 10

3-hydroxyisobutyrate dehydrogenase (EC 1.1.1.31) - Synechocystis sp. (strain PC)
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C>Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C:Accession: S76381
R:Kaneke, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S76381
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-290 <KAN>
A:Cross-references: EMBL:D64000; GB:AB001339; NID:g1001484; PIDN:BAAL0233.1; PID:g100160
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Superfamily: 3-hydroxyisobutyrate dehydrogenase; 3-hydroxyisobutyrate dehydrogenase ho
C:Keywords: oxidoreductase
F:6-267/Domain: 3-hydroxyisobutyrate dehydrogenase homology <HIB>

Query Match 11.8%; Score 338; DB 2; Length 290;
Best Local Similarity 29.6%; Pred. No. 3.8e-15;
Matches 85; Conservative 50; Mismatches 144; Indels 8; Gaps 3;
Qy 269 KIGFLGLMGSGIVSNLLKMGHTVTVMNRRTAEKCDLFIQEGARLG----RTPAEVVSTC 324
Db 5 KIAVFLGVGMSPMAQNLVNGVQTVGYNTLRPS--VQEAAGKAGVYVTSIAVAANA 62
Qy 325 DITFACVSDPAKADLVLPSPVQGIKRGPKCVYDMSTVDADVTVELAQVIVSRGGRFLE 384
Db 63 DIILTCVGEKDVQQLILGSGGIAEYAKPQALLIDCSITGKTAAYELATNLKQLGRFLD 122
Qy 385 APVSGNQQLSNDGMVLAAAGDRGLYEDCSSCFQAMGKTSFFFLGEVGNAAKMLIVNVQ 444
Db 123 APVTGVDVGAINGTLTIWVGDDISDFEALPVLKSGIKETVHCSPSGSGQAVKLCNQVLC 182
Qy 445 GSPMATIAEGLTLAHVTGQSQQTLIDILNOQLASIFLDQKCONILQGNFKPDPFYLKYIQ 504
Db 183 GIHAIAAAEAIQLSEQLGIAPELVITCGSAGSAGSALTNLAPKMSAADPAPGFVWKHL 242
Qy 505 KDLRLAIALGDVAVNHTPMMAAANEVYKRAKALDQSDNDMSAVYRAY 551
Db 243 KDLRLVREA--AENGPLPGVTLAESLFTSVQLLGGEDQGSQAIRAY 287

RESULT 11

AB3380
3-hydroxyisobutyrate dehydrogenase (EC 1.1.1.31) [imported] - Brucella melitensis (strain
C:Species: Brucella melitensis
C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002
C:Accession: AB3380
R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goldsman, E.; Seikov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AD3252; PMID:11756688
A:Accession: AB3380
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-291 <KUR>
A:Cross-references: GB:AE008917; PIDN:AAL52205.1; PID:g17982987; GSPDB:GN00190
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME11024
A:Map position: 1

C:Superfamily: 3-hydroxyisobutyrate dehydrogenase; 3-hydroxyisobutyrate dehydrogenase hor
C:Keywords: oxidoreductase

Query Match 11.7%; Score 335; DB 2; Length 291;
Best Local Similarity 27.3%; Pred. No. 6e-15;
Matches 79; Conservative 65; Mismatches 133; Indels 12; Gaps 5;
Qy 266 TDKIGFLGLMGSGIVSNL-LKMGHTVTVMNRRTAEKCDLFIQEGARLGRTPAEVSTC 323
Db 2 TNVAVFLGLGVGMYPMAHGLKARGGHDVTVMNRRTAAKAAKAEQFGGNYAATPADAVKN 61
Qy 324 CDITFACVSDPAKADLVLPSPVQGIKRGPKCVYDMSTVDADVTVELAQVIVSRGGRFL 383
Db 62 ADFVFCIGLNDLRLAIALGDVAVNHTPMMAAANEVYKRAKALDQSDNDMSAV 121
Qy 384 EAPVSGNQQLSNDGMVLAAAGDRGLYEDCSSCFQAMGKTSFFFLGEVGNAAKMLIVNV 438
Db 122 DAPVSGGQAGAENGVLTMVVGASEPVEFRAKPVTEAVARSVGLMGPVSGQLAKMNVQIC 181
Qy 439 IVNMVQGSFMATIAEGLTLAHVTGQSQQTLIDILNOQLASIFLDQKCONILQGNFKPDP 498
Db 182 IAGLVQG-----LABGHFGKXAGLDIEKJIAVISKGAAGSWMENSRSTMNQKGYDFG 236
Qy 499 YLKYIQKDLRLAIALGDVAVNHTPMMAAANEVYKRAKALDQSDNDMSAV 547
Db 237 AVDMRKDLGICLDEANRNGALLPVTALVDQFYKEVKIGGGRWDTSSL 285

RESULT 12

T34859
probable dehydrogenase - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C>Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 28-Jul-2000
C:Accession: T34859
R:Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, February 1999
A:Reference number: Z21559
A:Accession: T34859
A>Status: preliminary; translated from GB/EMBL/DDBI
A:Molecule type: DNA
A:Residues: 1-296 <OLI>
A:Cross-references: EMBL:AL035478; PIDN:CAB36613.1; GSPDB:GN00070; SCOEDB:SC2G5.26C
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SC2G5.26C
C:Superfamily: 3-hydroxyisobutyrate dehydrogenase; 3-hydroxyisobutyrate dehydrogenase hon
F:7-268/Domain: 3-hydroxyisobutyrate dehydrogenase homology <HIB>

Query Match 11.7%; Score 334; DB 2; Length 296;
Best Local Similarity 28.6%; Pred. No. 7.1e-15;
Matches 82; Conservative 53; Mismatches 142; Indels 10; Gaps 2;
Qy 269 KIGFLGLMGSGIVSNLLKMGHTVTVMNRRTAEKCDLFIQEGARLGRTPAEVSTC 328
Db 6 KVAVIGLGMSPMSNLKAGVQVTGFTLEQKDLRLAAAGGVAGSIAEVRDADVV 65
Qy 329 ACVSDPAKADLVLPSPVQGIKRGPKCVYDMSTVDADVTVELAQVIVSRGGRFLEAPVS 388
Db 66 TMVPASPOVEAIAYGPDGILENAKSGALLVDMSSITPQTSVDLAKAAKDKGIRVLDPVS 125
Qy 389 GNOQLSNDGMVLAAAGDRGLYEDCSSCFQAMGKTSFFFLGEVGN-----AAKMLIVNV 443
Db 126 GGEAGIAEVLIMVGGGEADFEAKPLEALGKTVILCGPHSGSGQTVKAAANQLIVANNI 185
Qy 444 QGSFMATIAEGLTLAHVTGQSQQTLIDILNOQLASIFLDQKCONILQGNFKPDPFYLKYI 503
Db 186 Q-----ACAEAVFLEKSGVDLKAALDVNLGSLAGSTVLRKKDNFLGRDFKPGFRIDUH 240
Qy 504 QKDLRLAIALGDVAVNHTPMMAAANEVYKRAKALDQSDNDMSAVYRAY 550
Db 241 HKDMGIVTDAARNVGAALPVGAVVAQLVASLRAQGGDLGDLHSAIRA 287

```
RESULT 13
AC1558
3-hydroxyisobutyrate dehydrogenase (B. subtilis YkwC protein) homolog lin1004 [imported]
C:Species: Listeria innocua
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AC1558
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker,
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AC1558
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-286 <GLA>
A:Cross-references: GB:AL592022; PIDN:CAC96235.1; PID:g16413463; GSPDB:GN00178
A:Experimental source: strain Clp11262
C:Genetics:
A:Gene: lin1004
C:Superfamily: 3-hydroxyisobutyrate dehydrogenase; 3-hydroxyisobutyrate dehydrogenase hc
Query Match 11.6%; Score 333; DB 2; Length 286;
Best Local Similarity 26.4%; Pred. No. 7.9e-15;
Matches 75; Conservative 61; Mismatches 148; Indels 0; Gaps 0;
QY 268 KKIGFLGLMGSGIVSNLLKMGHTVTVVNRRTAEKCDLFIQEGARLGRTPAEVSTCDIT 327
DB 2 EKIGFVGTVGMSSMAXHLEAGYEVHYTRTKAEALLSQGALWESDPGSLGAKVDIL 61
QY 328 FACVSDPKAAKDLVLGSPGVLOGIRPGKCYVDMSTVDADTVTELAQVIVSRGGRFLAPV 387
DB 62 ISMWGYPKDVEQLYLGENGFLDNKAGSVAIDMTTSSPALAKKIAEAGHEKGIGVLDAPV 121
QY 388 SGNQQLSNDGMVLVLAAGDRGLYEDCSSCFQAMGKTSFFLGEVGNAAKMMLIVNVQGSF 447
DB 122 SGDIGAKNGTLAIHWGNAEDVFLKVPFELIGSSVILQDAGSGQHTKMNQIAIASN 181
QY 448 MATIAEGLTLAHVTGQSQOTLLDILNQQLASIFLDQKCNILQGNFKPDFVKYIQKDL 507
DB 182 MGVTEAIIYABAAGLNPSRLVDSISGGAAGWSLTNLIPRVLDQDFSPGFFIKFIKDM 241
QY 508 RLAIALGDVNHPTPMAAAANEVYKRAKALDOSDNDSMAVYRAY 551
DB 242 GIAISEAKQMGLELPLTLAEQMYQTAEQGLSEBGTQALIKYY 285
RESULT 14
B83979
3-hydroxyisobutyrate dehydrogenase BH2634 [imported] - Bacillus halodurans (strain C-125
C:Species: Bacillus halodurans
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: B83979
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: B83979
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-299 <STO>
A:Cross-references: GB:AP001516; GB:BA000004; NID:g10175192; PIDN:BA06353.1; GSPDB:GN00
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH2634
C:Superfamily: 3-hydroxyisobutyrate dehydrogenase; 3-hydroxyisobutyrate dehydrogenase hc
Query Match 11.6%; Score 331.5; DB 2; Length 299;
Best Local Similarity 29.8%; Pred. No. 1.1e-14;
Matches 84; Conservative 47; Mismatches 150; Indels 1; Gaps 1;
```

```
QY 268 KKIGFLGLMGSGIVSNLLKMGHTVTVVNRRTAEKCDLFIQEGARLGRTPAEVSTCDIT 327
DB 3 QKGVFVGLGTMLGPMTHLVKLDKGFETYVTKSRSPBIEALQYGAIEVESYKELMETADIV 62
QY 328 FACVSDPKAAKDLVLGSPGVLOGIRPGKCYVDMSTVDADTVTELAQVIVSRGGRFLAPV 387
DB 63 MTCLELPELVTVYEGEDGHIAGLSQKILIDHSTVDRETNRVVAEQIKEGGGFLAPV 122
QY 388 SGNQQLSNDGMVLVLAAGDRGLYEDCSSCFQAMGKTSFFLGEVGNAAKMMLIVNVQGSF 447
DB 123 SGGPMGAKAGTILTMCGEADSPERSKEVLGAYGDIYVHVGPIGSGTVVKLANNNVGVH 182
QY 448 MATIAEGLTLAHVTGQSQOTLLDILNQQLASIFLDQKCNILQGNFKPDFVKYIQKDL 507
DB 183 QAVLGECLYFVEKAGVDPATAYEIKRSAGFSKSMESVDAILDRAFDPRFSINLLHKDI 242
QY 508 RLAIALGDVNHPTPMAAAANEVYKRAKALDOSDNDSMAVYRAY 549
DB 243 GLALKLGEOLGIPLEMEVKEERVAAKE-QYGHEDVSAIIR 283
RESULT 15
AE1200
3-hydroxyisobutyrate dehydrogenase (B. subtilis YkwC protein) homolog lmo1005 [imported]
C:Species: Listeria monocytogenes
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AE1200
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AE1200
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-286 <GLA>
A:Cross-references: GB:NC 003210; PIDN:CAC99083.1; PID:g16410407; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo1005
C:Superfamily: 3-hydroxyisobutyrate dehydrogenase; 3-hydroxyisobutyrate dehydrogenase ho
Query Match 11.5%; Score 331; DB 2; Length 286;
Best Local Similarity 27.1%; Pred. No. 1.1e-14;
Matches 77; Conservative 59; Mismatches 148; Indels 0; Gaps 0;
QY 268 KKIGFLGLMGSGIVSNLLKMGHTVTVVNRRTAEKCDLFIQEGARLGRTPAEVSTCDIT 327
DB 2 EKIGFVGTVGMSSMAXHLEAGYEVHYTRTKAEALLDQKALWVETPGSLANKVDIL 61
QY 328 FACVSDPKAAKDLVLGSPGVLOGIRPGKCYVDMSTVDADTVTELAQVIVSRGGRFLAPV 387
DB 62 ISMWGYPKDVEQLYLGENGFLDNKAGSVAIDMTTSSPALAKKIAEAGHEKGIGVLDAPV 121
QY 388 SGNQQLSNDGMVLVLAAGDRGLYEDCSSCFQAMGKTSFFLGEVGNAAKMMLIVNVQGSF 447
DB 122 SGDIGAKNGTLISINVGSSEDFLKVPIFDILGSSVILQDAGAGQHTKMNQIAIASN 181
QY 448 MATIAEGLTLAHVTGQSQOTLLDILNQQLASIFLDQKCNILQGNFKPDFVKYIQKDL 507
DB 182 MGVTEAIIYABAAGLNPSRLVDSISGGAAGWSLTNLIPRVLDQDFSPGFFIKFIKDM 241
QY 508 RLAIALGDVNHPTPMAAAANEVYKRAKALDOSDNDSMAVYRAY 551
DB 242 GIAISEAKQMGLELPLTLAEQMYQTAEQGLSEBGTQALIKYY 285
RESULT 16
G71912
probable 3-hydroxyacid dehydrogenase - Helicobacter pylori (strain J99)
```


C;Species: *Helicobacter pylori*
 A;Variety: strain J99
 C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 08-Oct-1999
 C;Accession: G71912
 R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
 ; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
 Nature 397, 176-180, 1999
 A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
 A;Reference number: A71800; MUID:99120557; PMID:9923682
 A;Accession: G71912
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-285 <ARN>
 A;Cross-references: GB:AE001491; GB:AE001439; NID:g4155127; PIDN:AAD06158.1; PID:g415512
 A;Experimental source: strain J99
 C;Genetics:
 A;Gene: jhp0585
 C;Superfamily: 3-hydroxyisobutyrate dehydrogenase; 3-hydroxyisobutyrate dehydrogenase ho
 F;3-266/Domain: 3-hydroxyisobutyrate dehydrogenase homology <HIB>
 Query Match 11.5%; Score 330; DB 2; Length 285;
 Best Local Similarity 32.7%; Pred. No. 1.2e-14;
 Matches 92; Conservative 46; Mismatches 137; Indels 6; Gaps 4;
 QY 269 KIGFLGLGMSGIVSNLLKMGHTVTVMNRTAEKCDLFIQEGARLGRTPAEVSTCDITF 328
 DB 2 KIGWIGLGQGLPMVTRLLDGGIEVGYNRSPTAPISAKGAKVYGNATLVRDYPVIF 61
 QY 329 ACVSDPKAAKDLVLGSPGVLOGIRPGKCYVDMSTVDADTVELAQVIVSRGRFLEAPVS 388
 DB 62 TMLUSD-KAALDAVLAPKFWQMSK-KIVVMNTIAPLESLSLEKTAQKHQATYLEAPVS 118
 QY 389 GNOQLSNDGMLVILAAGDRGLYEDCSSCFQAMGKTSFPLGEVGNAAKMLIVNMVQGSFM 448
 DB 119 GSVGAPATNGTLLILFGGSEAVLNPLQKIFSLVGKTKTFHFGDVGKSGAKLVLSLLGIFG 178
 QY 449 ATTAEGTLTAHVTCGSOQTLLDILNQGLASIFLDQKCONILQGNFKPDPFLVKYIOKDLR 508
 DB 179 VAYSEALLAKRLGVDAESFLQIIGSGMNSPLFQAKKGWMLQDNYPAAFSLKLMKAIIR 238
 QY 509 LA-IALGDVAVNHPPTMAAAANEVYKRAKALDQSDNDMSAVY 548
 DB 239 LAKNEAGEAMK--LPFLFQAEELYSAEKLSGGLDMAVY 277
 RESULT 17
 B81065
 3-Hydroxyacid dehydrogenase NM1584 [imported] - Neisseria meningitidis (strain MC58 ser
 C;Species: *Neisseria meningitidis*
 C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Nov-2001
 C;Accession: B81065
 R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
 Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
 Li, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.
 Science 287, 1809-1815, 2000
 A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
 A;Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.
 A;Reference number: A81000; MUID:20175755; PMID:10710307
 A;Accession: B81065
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-289 <TET>
 A;Cross-references: GB:AE002509; GB:AE002098; NID:g7226832; PIDN:AAF41937.1; PID:g722683
 A;Experimental source: serogroup B, strain MC58
 C;Genetics:
 A;Gene: NM1584
 C;Superfamily: 3-hydroxyisobutyrate dehydrogenase; 3-hydroxyisobutyrate dehydrogenase ho
 F;9-266/Domain: 3-hydroxyisobutyrate dehydrogenase homology <HIB>
 Query Match 11.5%; Score 329; DB 2; Length 289;
 Best Local Similarity 30.0%; Pred. No. 1.5e-14;
 Matches 84; Conservative 52; Mismatches 140; Indels 4; Gaps 2;
 QY 269 KIGFLGLGMSGIVSNLLKMGHTVTVMNRTAEKCDLFIQEGARLGRTPAEVSTCDITF 328
 DB 8 QIGWIGLGQGLPMVTRLLDGGIEVGYNRSPTAPISAKGAKVYGNATLVRDYPVIF 67
 QY 329 ACVSDPKAAKDLVLGSPGVLOGIRPGKCYVDMSTVDADTVELAQVIVSRGRFLEAPVS 388
 DB 68 LMVSDYAAVCDIL---NGVRDGL-AGKIIVMNSTISPTENLAVKALVEAAGGQFAEAPVS 123
 QY 389 GNOQLSNDGMLVILAAGDRGLYEDCSSCFQAMGKTSFPLGEVGNAAKMLIVNMVQGSFM 448
 DB 124 GSVGAPATNGTLLILFGGSEAVLNPLQKIFSLVGKTKTFHFGDVGKSGAKLVLSLLGIFG 183
 QY 449 ATTAEGTLTAHVTCGSOQTLLDILNQGLASIFLDQKCONILQGNFKPDPFLVKYIOKDLR 508
 DB 184 EAYSEAMLARQFGIDTDTIVEAIGGSAMDSPMFQTKSLWANREFPPAFALKHASKDLN 243
 QY 509 LAIALGDVAVNHPPTMAAAANEVYKRAKALDQSDNDMSAVY 548
 DB 244 LAVKELEQAGNTLPAVETVAASRKAVEAGYGEQDVSGVY 283
 RESULT 19
 F83447
 probable 3-hydroxyisobutyrate dehydrogenase PA1576 [imported] - *Pseudomonas aeruginosa* (s
 C;Species: *Pseudomonas aeruginosa*
 C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

QY 269 KIGFLGLGMSGIVSNLLKMGHTVTVMNRTAEKCDLFIQEGARLGRTPAEVSTCDITF 328
 DB 8 QIGWIGLGQGLPMVTRLLDGGIEVGYNRSPTAPISAKGAKVYGNATLVRDYPVIF 67
 QY 329 ACVSDPKAAKDLVLGSPGVLOGIRPGKCYVDMSTVDADTVELAQVIVSRGRFLEAPVS 388
 DB 68 LMVSDYAAVCDIL---NGVRDGL-AGKIIVMNSTISPTENLAVKALVEAAGGQFAEAPVS 123
 QY 389 GNOQLSNDGMLVILAAGDRGLYEDCSSCFQAMGKTSFPLGEVGNAAKMLIVNMVQGSFM 448
 DB 124 GSVGAPATNGTLLILFGGSEAVLNPLQKIFSLVGKTKTFHFGDVGKSGAKLVLSLLGIFG 183
 QY 449 ATTAEGTLTAHVTCGSOQTLLDILNQGLASIFLDQKCONILQGNFKPDPFLVKYIOKDLR 508
 DB 184 EAYSEAMLARQFGIDTDTIVEAIGGSAMDSPMFQTKSLWANREFPPAFALKHASKDLN 243
 QY 509 LAIALGDVAVNHPPTMAAAANEVYKRAKALDQSDNDMSAVY 548
 DB 244 LAVKELEQAGNTLPAVETVAASRKAVEAGYGEQDVSGVY 283
 RESULT 18
 E81802
 hypothetical protein NMA1773 [imported] - *Neisseria meningitidis* (strain Z2491 serogroup
 C;Species: *Neisseria meningitidis*
 C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Nov-2001
 C;Accession: E81802
 R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morell
 ; Holroyd, S.; Jørgels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
 Nature 404, 502-506, 2000
 A;Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.
 A;Reference number: A81775; MUID:20222556; PMID:10761919
 A;Accession: E81802
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-289 <PAR>
 A;Cross-references: GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CAB85001.1; PID:g7380415
 A;Experimental source: serogroup A, strain Z2491
 C;Genetics:
 A;Gene: NMA1773
 C;Superfamily: 3-hydroxyisobutyrate dehydrogenase; 3-hydroxyisobutyrate dehydrogenase hon
 F;9-266/Domain: 3-hydroxyisobutyrate dehydrogenase homology <HIB>
 Query Match 11.5%; Score 329; DB 2; Length 289;
 Best Local Similarity 30.0%; Pred. No. 1.5e-14;
 Matches 84; Conservative 52; Mismatches 140; Indels 4; Gaps 2;
 QY 269 KIGFLGLGMSGIVSNLLKMGHTVTVMNRTAEKCDLFIQEGARLGRTPAEVSTCDITF 328
 DB 8 QIGWIGLGQGLPMVTRLLDGGIEVGYNRSPTAPISAKGAKVYGNATLVRDYPVIF 67
 QY 329 ACVSDPKAAKDLVLGSPGVLOGIRPGKCYVDMSTVDADTVELAQVIVSRGRFLEAPVS 388
 DB 68 LMVSDYAAVCDIL---NGVRDGL-AGKIIVMNSTISPTENLAVKALVEAAGGQFAEAPVS 123
 QY 389 GNOQLSNDGMLVILAAGDRGLYEDCSSCFQAMGKTSFPLGEVGNAAKMLIVNMVQGSFM 448
 DB 124 GSVGAPATNGTLLILFGGSEAVLNPLQKIFSLVGKTKTFHFGDVGKSGAKLVLSLLGIFG 183
 QY 449 ATTAEGTLTAHVTCGSOQTLLDILNQGLASIFLDQKCONILQGNFKPDPFLVKYIOKDLR 508
 DB 184 EAYSEAMLARQFGIDTDTIVEAIGGSAMDSPMFQTKSLWANREFPPAFALKHASKDLN 243
 QY 509 LAIALGDVAVNHPPTMAAAANEVYKRAKALDQSDNDMSAVY 548
 DB 244 LAVKELEQAGNTLPAVETVAASRKAVEAGYGEQDVSGVY 283
 RESULT 19
 F83447
 probable 3-hydroxyisobutyrate dehydrogenase PA1576 [imported] - *Pseudomonas aeruginosa* (s
 C;Species: *Pseudomonas aeruginosa*
 C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: A98338
A;Status: preliminary
A:Molecule type: DNA
A;Residues: 1-315 <KUR>
A;Cross-references: GB:AEO07870; PIDN:AAK90227.1; PID:gI5160240; GSPDB:GN00170
C;Genetics:
A;Gene: AGR_L_3303
A;Map position: linear chromosome
C;Superfamily: 3-hydroxyisobutyrate dehydrogenase; 3-hydroxyisobutyrate dehydrogenase hc

Query Match 11.1%; Score 318; DB 2; Length 315;
Best Local Similarity 27.0%; Pred. No. 9e-14;
Matches 75; Conservative 58; Mismatches 139; Indels 6; Gaps 2;

QY 268 KKIGFLGLGMSGIVSNLLKMGHVTVWNRTAEKCDLFIQEGARLGRTPAEVSTCIDT 327
Db :
12 RDIAVI GTGIMGAPMAARLA EAGFAVRWNRRRAEKAAILAEKGVRQAATTIDA AEADV V 71

QY 328 FACVS DPKAAKDVLGPGSVLG IGPCKYVDMS TADVTVELAQVIVSRGRFLEAPV 387
Db :
72 ICM LSGPVCDEV L LGPGSVSAMKPSGVLLVMSSIPVDSAREQAARA HGV RVVDAPV 131

QY 388 SGNQQLSNDGMVLILAADRGLEYDCSCFOAMGKTFFLGEVGNAAKOMLVINMVQGSF 447
Db :
132 SGGEKAIEGT LAIMAGEGEQRDVALRPLNLCLGRVT-HVGFVCGSLAKLANQLIVAST 190

QY 448 MATIABGLTLAHVTGQSQQTLDDLNGQLASIFLDQKCMILOGNFDPFLKYIQKDL 507
Db :
191 IC AVAEAL TLVEAGGDPAQVRQALLGGFAESTVFROHGKRMVEGDFRP GGPAKYQVKDT 250

QY 508 RLAI ALGDVANHPTPM AAANEVYKR-----AKALDQS 540
Db :

RESULT 23
AI2944
3-hydroxyisobutyrate dehydrogenase mmsB [imported] - Agrobacterium tumefaciens (strain C)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C;Accession: AI2944
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.; erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan; ; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AI2944
A;Status: preliminary
A:Molecule type: DNA
A;Residues: 1-315 <KUR>
A;Cross-references: GB:AEO08689; PIDN:AAL43975.1; PID:gI7741531; GSPDB:GN00187
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: mmsB
A;Map position: linear chromosome
C;Superfamily: 3-hydroxyisobutyrate dehydrogenase; 3-hydroxyisobutyrate dehydrogenase hc

Query Match 11.1%; Score 318; DB 2; Length 315;
Best Local Similarity 27.0%; Pred. No. 9e-14;
Matches 75; Conservative 58; Mismatches 139; Indels 6; Gaps 2;

QY 268 KKIGFLGLGMSGIVSNLLKMGHVTVWNRTAEKCDLFIQEGARLGRTPAEVSTCIDT 327
Db :
12 RDIAVI GTGIMGAPMAARLA EAGFAVRWNRRRAEKAAILAEKGVRQAATTIDA AEADV V 71

QY 328 FACVS DPKAAKDVLGPGSVLG IGPCKYVDMS TADVTVELAQVIVSRGRFLEAPV 387
Db :
72 ICM LSGPVCDEV L LGPGSVSAMKPSGVLLVMSSIPVDSAREQAARA HGV RVVDAPV 131

C;Genetics:
A;Map: Arsp:Tl3K14.90
A;Gene position: 4
A;Introns: 38/3; 68/2; 106/3; 157/1; 208/3; 240/3; 288/2; 316/3; 343/2
C;Superfamily: 3-hydroxyisobutyrate dehydrogenase; 3-hydroxyisobutyrate dehydrogenase hc
F;40-344/Domain: 3-hydroxyisobutyrate dehydrogenase homology <H1B>

Query Match 11.0%; Score 316; DB 2; Length 371;
Best Local Similarity 25.3%; Pred. No. 1.5e-13;
Matches 88; Conservative 58; Mismatches 154; Indels 48; Gaps 6;

Qy 248 STSIQAOSTAVNGSITPTDKKIGFLGILMGSGIYSNLLKMGHVTVMNRTEKCDLEI 307
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 21 SGSLHRFSNSQN---QFNVGFTGLNMGRFMVNNLIRAGKYKVTHDINRDMKMT 77
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

Qy 308 QEGARLGRTPAEVSVCSDITTFACVPDPKAADVLVPGSVLQG---LRPKCYVDMSIVD 364
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 78 EMGVSSRPPEYVAQDSSEVIIMLPSSSHVMDVVTGTNGLLGENDIRPA-LFIDSSTID 136
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

Qy 365 ADTVTELAAQIVSRGR-----FLRAPVSGN 390
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 137 PQTTRKISLAIVENCNLKEKRASFCSCQEVEIVSFSLVSSIFDVDNWKEPWLDAVSGG 196
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

Qy 391 QQLSNDGMVLIIAAGRGLGYEDCGSCCFQAMGKTSSFPLGEVGNAAKMMLIVNMVGGSFAT 450
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 197 VLAAEAGTLTFMVGPPEDAYLAARPILQSMGRTSIYC GGSGNGSAAKICNNLAMAVSMLG 256
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

Qy 451 IAEGLTIAHTVGSQQOTLLDIIN--QGOLASIFLDOKCONIQG-----NEKPFPLYKVI 503
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 257 TSEALALGQSLGISASTLTLELVNTSSRCRWSSDAYNPVPGVMKGVPSSRDYNGGFASKLM 316
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

Qy 504 QKDRLRAIALGDVANNHPTMAAAANEVYKRAKALDQSDNDMSAVYRAY 551
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 517 AFDNIANAASAFVGHKSPTISKAOEIVYKKMCCEGHETKDFSCVFRRHF 564
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

RESULT 26
C83232
probable 3-hydroxyisobutyrate dehydrogenase PA3312 [imported] - Pseudomonas aeruginosa
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: C83232
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: C83232
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-296 <STO>
A:Cross-references: GB:AE004753; GB:AE004091; NID:g9949433; PIDN:AG06700.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA3312
A:Enzyme: 3-hydroxyisobutyrate dehydrogenase: 1-hydroxyisobutyrate dehydrogenase h

	Query Match	10.8%;	Score 310;	DB 2;	Length 296;
	Best Local Similarity	28.6%;	Pred. No. 2.8e-13;		
	Matches 83;	Conservative 57;	Mismatches 136;	Indels 14;	Gaps 6;
Qy	270	IGFGLGIMSGSYNSLLKMGHTVTVMNRTASKDCLFIQEGARLGRTPAEVSTCDITFA	329		
Db	8	LAFAGLGMGVMPCRLLAAGYPLAVWNRSPGKRELLAAGAKAVEVPAELAADAELML	67		
Qy	330	CVSPDKAAKDLVLGPGSVGLQIRPGCVYDMSTVDADTVTELAQVTVSR-GRFLEAPVS	388		
Db	68	CLADTAAVREVVFGAGGIVENARPQGLLYDFPSAEPAATREMAAAELEARGCVRWVDAVS	127		
Qy	389	GNQQLSNDGMLVILAAGRDRLYEDCSSCFQAMGKTSFFFLGCEVGNAAKMYLIVMWQGSFM	448		
Db	128	GGTCPAEGSSLAIMGAGRAADTERLPVLSLGRLTRMGVEVG-AGQVTKVCNQ-----M	181		

```

449  ATTAEGTLTAHVGGQSQTLDDI-INQOGLASTFLDKQCNTLL-----QGKPKP-DFYLLK 501
      ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: :::
Db    182  IVACNALVTIAEVVALAAGVADSLVAPALAGGAFADSKPLQILAPQMAASRPFVQKHVR 241
      ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: :::
Qy    502  YIQKDLRLAIALGDVANHPTFMAAANEVYKRAKALDQSDNDSAVYRAY 551
      ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: :::
Db    242  TLLKOLDTAVKLSREQGAATPMSGLAAQLMRLHGSQGYLDRDPATLVEQY 291
      ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: :::

RESULT 27
B96015
probably 2-hydroxy-3-oxopropionate reductase (BC 1.1.1.1.60) [imported] - Sinorhizobium meliloti
C/Species: Sinorhizobium meliloti
C/Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 14-Sep-2001
C/Accession: B96015
R/Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F. J. J. Proc. Natl. Acad. Sci. U.S.A. 98, 9899-9894, 2001
A/Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-fi-
A/Reference number: A95842; MUID:21396508; PMID:11481431
A/Accession: B96015
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-294 <KUR>
A/Cross-references: GB:AL591985; PIDN:CAC49786.1; PID:g15141273; GSPDB:GN00167
A/Experimental source: strain 1021, megaplasmid pSymB
R/Galibert, F.; Finan, T.M.; Long, S.K.; Puhler, A.; Abola, P.; Ampe, F.; Baril-
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fish-
ler, H.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A/Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.;
hehault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.
A/Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A/Reference number: A96039; MUID:21368234; PMID:11474104

```

[illegible]

RESULT 28
G96013
Conserved hypothetical protein SMB20668 [imported] - *Sinorhizobium meliloti* (S)
C:Species: *Sinorhizobium meliloti*
C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C/Accession: G96013
C/Finan: T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.

Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
 A:Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-fixing endo
 A:Reference number: A95842; MUID:21396508; PMID:11481431
 A:Accession: G96013
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-301 <KUR>
 A:Cross-references: GB:AL591985; PIDN:CAC49775.1; PID:g151141262; GSPDB:GN00167
 A:Experimental source: strain 1021, megaplasmid pSymB
 R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
 P.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A:Authors: Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
 Hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
 A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 A:Contents: annotation
 C:Genetics:
 A:Gene: SMB20668
 A:Genome: plasmid
 C:Superfamily: 3-hydroxyisobutyrate dehydrogenase; 3-hydroxyisobutyrate dehydrogenase ho
 Query Match 10.6%; Score 303.5; DB 2; Length 301;
 Best Local Similarity 26.4%; Pred. No. 7.7e-13;
 Matches 75; Conservative 56; Mismatches 152; Indels 1; Gaps 1;
 QY 269 KIGFLGLGMSGIVSNLLKMGHTVWNRATAEKCDLFTQEGARLGRTPAEVWSTCDITF 328
 DB 4 KIAVLGLSGMFGMACSKMSAGLDVLGVDVAPPAVERFVAEGRGAGTPGAVTGADIIV 63
 QY 329 ACVSDPKAAKDLVLGSPGVLOGIRPKGKCYVDMSTVDADTVTELAQVIVSRGRFLAPVS 388
 DB 64 SIIVSGAQTEAVLFGPNVGAGMKPGAAFISSATMDPAIRDLAQLRLEALGHYLDAPIS 123
 QY 389 GNOQLSNDGMLVLAAGDRGLYEDCSSCFQAMGKTSFFL-GEVGNAAKMMLIVNVWGSPF 447
 DB 124 GGAARAKGELTIVASGSPQAFARPAALDAAKAVVELGCTAGTGAAFKMINQLLAGVH 183
 QY 448 MATIAEGLTLAHVTGQSQOQLDILNQGLASIFLDQKCONILQGNFKPDFYLYKIQKDL 507
 DB 184 IAAACEALAPAAKGLDLKVEVITASAGNSWFFENRIPHLVAGDVAPLSAIEIFVKDL 243
 QY 508 RLALAGDVAHVHPTMAAANEVYKRAKALDQSDNMSAVYRAY 551
 DB 244 GIVQDMARERYPPVPLVAAALQMTYLAASGAGMGRRDDSSLARLY 287
 RESULT 29
 D64782
 3-hydroxyisobutyrate dehydrogenase homolog b0509 - Escherichia coli (strain K-12)
 C:Species: Escherichia coli
 C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
 C:Accession: D64782
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
 A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of Escherichia coli K-12.
 A:Reference number: A64720; MUID:97426617; PMID:9278503
 A:Accession: D64782
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-292 <BLAT>
 A:Cross-references: GB:AE000157; GB:U00096; NID:g1786716; PIDN:AAC73611.1; PID:g1786719;
 A:Experimental source: strain K-12, substrain MGL655
 C:Superfamily: 3-hydroxyisobutyrate dehydrogenase; 3-hydroxyisobutyrate dehydrogenase ho
 F; 3-263/Domain: 3-hydroxyisobutyrate dehydrogenase homology <HIS>
 Query Match 10.5%; Score 301.5; DB 2; Length 292;
 Best Local Similarity 27.7%; Pred. No. 1e-12;
 Matches 76; Conservative 51; Mismatches 152; Indels 1; Gaps 1;
 QY 269 KIGFLGLGMSGIVSNLLKMGHTVWNRATAEKCDLFTQEGARLGRTPAEVWSTCDITF 328

DB 2 KLGLFGLGIMGTPMAINLARAGHQLHV-TTIGPVADELLSLGAVSVETARQVTHSDIIF 60
 QY 329 ACVSDPKAAKDLVLGSPGVLOGIRPKGKCYVDMSTVDADTVTELAQVIVSRGRFLAPVS 388
 DB 61 INVPTDTPQVEEVLFGENGCTKASLAKTKTIVDMSSISPIETKRFARQVNLGGDYLDPVS 120
 QY 389 GNOQLSNDGMLVLAAGDRGLYEDCSSCFQAMGKTSFFLGEVGNAAKMMLIVNVWGSPF 448
 DB 121 GGEIGARETSLIMVGGDEAVFERVKPFLFELGKNITLVGGNGDGTCKVANQIIVALMI 180
 QY 449 ATIAEGLTLAHVTGQSQOQLDILNQGLASIFLDQKCONILQGNFKPDFYLYKIQKDLR 508
 DB 181 EAVSEALLFASKAGADPVVRQALMGSGFASSRILEVHGEMIKRTFNFQFKIALHOKDLN 240
 QY 509 LAIALGDVAHVHPTMAAANEVYKRAKALDQSDNMSAVYRA 550
 DB 241 LALQSAKALALNLTATCOELFNTCAANGGSQLDHSALVQA 282

RESULT 30

E96736

Probable dehydrogenase F23N20.17 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001

C:Accession: E96736

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 ansen, N.F.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 Nature 408, 816-820, 2000A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.
 Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: E96736

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-297 <STO>

A:Cross-references: GB:AE005173; NID:G6714326; PIDN:AAF26019.1; GSPDB:GN00141

C:Genetics:

A:Gene: F23N20.17

A:Map position: 1

C:Superfamily: 3-hydroxyisobutyrate dehydrogenase; 3-hydroxyisobutyrate dehydrogenase ho

Query Match

Best Local Similarity 10.5%; Score 301.5; DB 2; Length 297;

Matches 76; Conservative 56; Mismatches 130; Indels 7; Gaps 2;

QY 263 ITPDCKKIFGLGMSGGVSNLLKMGHTVWNRATAEKCDLFTQEGARLGRTPAEVVS 322

DB 9 IDPSKTRIGIGIMGSAMVSHIIAAGYSVVVARDLRKTDQLQTKGARIANSFKELAE 68

QY 323 TCIDTEACVSDPKAAKDLVLGSPGVLOGIRPKGKCYVDMSTVDADTVTELAQVIVSRGRF 382

DB 69 MSDVFTVFNVDNDRSLLDGGLTPGGVTDVMTSSKPLAREIHAEARRNCWA 128

QY 383 LEAPVSGNOQLSNDGMLVLAAGDRGLYEDCSSCFQAMGKTSFFLGEVGNAAKMMLIVNM 442

DB 129 VDAVSGDGAGAREGTGIFAGGSEIVEWLSPPMKNG-ITVYMGAGGSGQSKGNQI 187

QY 443 VQGSFWATIAEGLTLAHVTGQSQOQLDILNQGLASIFLDQKCONILQGNFKPDFYLYK 502

DB 188 AGASNLVGLAEGIVFAEKAGLDTVKLEAVKDGAGSAVNRFLGEMIVKEDYRATGAEY 247

QY 503 IQKDLRLALATAGDVAHVHPTMAAANEVY 531

DB 248 MVKD-----LGVAAABAAEPGAALSKQLF 270

Search completed: September 16, 2004, 07:31:11

us-10-067-482-2.rpr

Thu Sep 16 09:36:52 2004

Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2004, 07:28:53 ; Search time 13 Seconds
(without alignments)
2214.984 Million cell updates/sec

Title: US-10-067-482-2

Perfect score: 2866

Sequence: 1 MAAVSLRLGLDVLWGKLGKRYP.....AKALDQSDNDMSAVRAYIH 553

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	368	12.8	288	1	YKWC_BACSU
2	367	12.8	294	1	GARR_ECOLI
3	338	11.8	290	1	Y229_SYNY3
4	328	11.4	347	1	D3HI_ARATH
5	301.5	10.5	292	1	Q5UC0_ARABIDOPSY
6	297	10.4	335	1	GLXR_ECOLI
7	290	10.1	335	1	DH3I_MOUSE
8	289.5	10.1	298	1	D3HI_RAT
9	287.5	10.0	298	1	YIHU_SALTY
10	281	9.8	336	1	YIHU_ECOLI
11	278.5	9.7	324	1	D3HI_HUMAN
12	261.5	9.1	298	1	D3HI_DROME
13	259.5	9.1	299	1	MMSB_PSEAE
14	239	8.3	302	1	DH3I_CAREL
15	234.5	8.2	284	1	YGBJ_ECOLI
16	224.5	7.8	261	1	MMSB_MYCTU
17	209.5	7.3	295	1	YFJR_BACSU
18	198	6.9	301	1	YGBJ_HAEIN
19	186.5	6.5	237	1	HDGF_MOUSE
20	185.5	6.5	237	1	HDGF_MOUSE
21	185.5	6.5	240	1	YCY3_SHEFR
22	148.5	5.2	1360	1	MSH6_HUMAN
23	146	5.1	468	1	6PGD_ECOLI
24	145.5	5.1	1358	1	MSH6_MOUSE
25	143.5	5.0	485	1	6PGD_CUNEL
26	141	4.9	468	1	6PGD_CUNEL
27	138.5	4.8	492	1	6PGD_SCHPO
28	137	4.8	468	1	6PGD_SHIFL
29	136	4.7	468	1	6PGD_ECOLI
30	135.5	4.7	1427	1	MES4_DROME
31	135.5	4.7	2459	1	MAPE_RAT
32	132.5	4.6	470	1	6PGD_SYNP7
33	132.5	4.6	484	1	6PGD_HAEIN

34	132	4.6	468	1	6PGD_SALTY
35	129.5	4.5	468	1	6PGD_BACSU
36	129	4.5	445	1	6PGD_KLETE
37	129	4.5	598	1	CYL1_HUMAN
38	128.5	4.5	468	1	6PGD_BUCBP
39	127.5	4.4	517	1	6PGD_CANAL
40	127.5	4.4	1359	1	ATRX_CAREL
41	127	4.4	445	1	6PGD_CITAM
42	127	4.4	2567	1	MI8B_HUMAN
43	125.5	4.4	479	1	6PGD_TRYBB
44	123.5	4.3	468	1	6PG2_BACSU
45	123	4.3	445	1	6PGD_CITFR
46	123	4.3	467	1	6PGD_BACLI
47	123	4.3	468	1	6PGD_STAN
48	122.5	4.3	908	1	DM3A_MOUSE
49	122	4.3	468	1	6PGD_STAAM
50	121.5	4.2	2144	1	GLT1_YEAST
51	120.5	4.2	488	1	6PGD_TREPA
52	120.5	4.2	909	1	DM3A_HUMAN
53	120	4.2	1186	1	PKCB_HUMAN
54	119.5	4.2	741	1	ERF2_FICPI
55	118.5	4.1	2464	1	MAPB_MOUSE
56	118	4.1	445	1	6PGD_ESCVU
57	118	4.1	468	1	6PGD_STAP
58	118	4.1	634	1	KNOB_PLAFG
59	117	4.1	445	1	6PGD_KLEPL
60	117	4.1	2696	1	NSD1_HUMAN
61	115	4.0	977	1	REMF_HUMAN
62	115	4.0	1402	1	Y232_HUMAN
63	114.5	4.0	472	1	6PGD_LACLA
64	114.5	4.0	959	1	IF2_BRUME
65	114.5	4.0	959	1	IF2_BRUSE
66	114	4.0	445	1	6PGD_CITDI
67	114	4.0	445	1	6PGD_SHIBO
68	114	4.0	445	1	6PGD_SHIDY
69	114	4.0	445	1	6PGD_SHISO
70	113.5	4.0	492	1	6PG2_YEAST
71	113	3.9	829	1	IF2_HAEIN
72	113	3.9	2332	1	PCLO_FMDVA
73	112.5	3.9	5085	1	PCLO_RAT
74	112	3.9	484	1	6PGD_ACTAC
75	112	3.9	494	1	SFR4_HUMAN
76	111	3.9	562	1	BS69_HUMAN
77	111	3.9	743	1	FTSX_CLOTE
78	111	3.9	1210	1	AF4_HUMAN
79	111	3.9	1362	1	BRD4_HUMAN
80	110.5	3.9	506	1	CDYM_HUMAN
81	110.5	3.9	543	1	TLPI_MOUSE
82	110.5	3.9	633	1	MLH_TEETH
83	110.5	3.9	1581	1	PPRE_HUMAN
84	110	3.8	875	1	ARS2_MOUSE
85	109.5	3.8	472	1	6PGD_LALCL
86	109.5	3.8	494	1	SFR4_MOUSE
87	109.5	3.8	1007	1	PR4B_HUMAN
88	109	3.8	700	1	TRDN_CANFA
89	109	3.8	907	1	IF2_VIBVU
90	109	3.8	907	1	IF2_VIBVY
91	109	3.8	1101	1	PHF2_HUMAN
92	109	3.8	1462	1	NKCR_HUMAN
93	109	3.8	2492	1	ATRX_PANTR
94	108.5	3.8	690	1	RHO_MICLU
95	108	3.8	481	1	6PGD_CERCA
96	108	3.8	611	1	SAH3_HUMAN
97	107.5	3.8	494	1	SFR4_RAT
98	107.5	3.8	705	1	TRDN_RABIT
99	107.5	3.8	780	1	CDL2_HUMAN
100	107.5	3.8	892	1	IF2_YERPE

ALIGNMENTS

RESULT 1

YKWC BACSU
ID YKWC BACSU STANDARD; PRT; 288 AA.
AC O34948;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical oxidoreductase YKWC (EC 1.1.-.-).
GN YKWC OR BSU13960.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RC STRAIN=168;
RT Scanlan E., Devine K.M.;
RT "Sequence of the Bacillus subtilis chromosome from ykua to cse-15.";
RT Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RC MEDLINE=98044033; PubMed=9384377;
RA Kunst P., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azavedo F., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrazi E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Chim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaere-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinot S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scofield F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzneger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256(1997).
CC -!- SIMILARITY: Belongs to the 3-hydroxyisobutyrate dehydrogenase
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ222587; CRA10859.1; -
CC EMBL; Z99111; CAB13269.1; -
CC EIR; B69870; B69870.
CC Subtilist; BG13328; ykwc.
CC InterPro; IPR002204; 3hydroxisobut_dh.
CC InterPro; IPR006183; 6PGD.
CC InterPro; IPR006115; 6PGD_NAD.
CC Pfam; PF03446; NAD_binding_2; 1.
CC PRINTS; PR00076; 6PGDHRGNASE.
CC PROSITE; PS00895; 3_HYDROXYISOBUT_DH; 1.
KW Hypothetical protein; Oxidoreductase; NAD; Complete proteome.
FT ACT SITE 172 BY SIMILARITY.
SQ SEQUENCE 288 AA; 30711 MW; 976DD098DBA7A30 CRC64;
Query Match 12.8%; Score 368; DB 1; Length 288;
Best Local Similarity 26.7%; Pred. No. 3.2e-16;
Matches 76; Conservative 65; Mismatches 144; Indels 0; Gaps 0;
QY 268 KKIGFLGGLMGSGIVSNLLKMGHTVTVNNTAKKCDLFIQEGARLGRTPAEVSTCDIT 327
DB 3 KTIQFGLGVGKMSASHILNDGHPVLVYTKKAESILQKGAIKWIKTDVSKDSKADVI 62
QY 328 FACVSDPKAAKDLVLGPGSVLQIRPGKCYVDMSTVDADTVELAQVIVSGRFLAPV 387
DB 63 ITWGYPSDVEVYFGSNGIENAKGAYLIDMTTSKPSLAKTAAEAKEKALFALDAPV 122
QY 388 SGNQQLSNDGMLVILAGDRGLYEDCCSCFOAMGKTSFFELGCVGNAKMLVNNVQGSF 447
DB 123 SGGDIGAQNGTLAIWVGGEKAEACMPFISLMGENIYQGPAGSGQHTKMCNQIAIAG 182
QY 448 MATTAEGLTLAHVTGQSQQTLTLLIINQGLASIFLDQKQNLQCNFKPDPFVLYKIQKDL 507
DB 193 MIGVAEMAYAKSKLEPENVLKSTTTGAAGSWSLSNLAPRMLQGNFEPGFVYKFIKDM 242
QY 508 RLATAGDAVNHPTMAAANEVYKRAKALDQSDNDMSAVYRAYI 552
DB 243 GIALEEALMGEMPGLSLAKSLYDKLAAQGEENSGTQSIYKLWV 287
RESULT 2
ID GARR_ECOLI STANDARD; PRT; 294 AA.
AC P23523;
DT 01-NOV-1991 (Rel. 20, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 2-hydroxy-3-oxopropionate reductase (EC 1.1.1.60) (Tartronate
DE semialdehyde reductase) (TSAR).
GN GARR OR B3125 OR C3880.
OS Escherichia coli and
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 217992;
RN [1] SEQUENCE FROM N.A.
RP STRAIN=K12;
RC MEDLINE=91154140; PubMed=1705543;
RA Komine Y., Inokuchi H.;
RA "Precise mapping of the rnpB gene encoding the RNA component of RNase
RT P in Escherichia coli K-12.";
RL J. Bacteriol. 173:1813-1816(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / CFT073 / ATCC 700928;
RC MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glaesner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=O6:HL / CFT073 / ATCC 700928;
RC MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli".
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).

```

RN CHARACTERIZATION.
RX MEDLINE=9847507; PubMed=9772162;
RA Hubbard B.K., Koch M., Palmer D.R., Babbitt P.C., Gerlt J.A.;
RT "Evolution of enzymatic activities in the enolase superfamily:
PT characterization of the (D)-glucarate/galactarate catabolic pathway
RT in Escherichia coli.";
RL Biochemistry 37:14369-14375(1998).
RN GENE NAME.
RX MEDLINE=20225875; PubMed=10762278;
RA Monterrubio R., Baldoma L., Obradors N., Aguilar J., Badia J.;
RT "A common regulator for the operons encoding the enzymes involved in
RT D-galactarate, D-glucarate, and D-glycerate utilization in
RT Escherichia coli.";
RL J. Bacteriol. 182:2672-2674(2000).
CC -1- CATALYTIC ACTIVITY: (R)-glycerate + NAD(P)(+) = 2-hydroxy-3-
CC oxopropanoate + NAD(P)H.
CC -1- PATHWAY: D-galactarate metabolism; third step.
CC -1- SIMILARITY: Belongs to the 3-hydroxyisobutyrate dehydrogenase
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D90212; BAA14238.1; ALT INIT.
DR EMBL; U18997; AAA57928.1; ALT INIT.
DR EMBL; AE000394; AAC76159.1; ALT INIT.
DR EMBL; AE016767; AAN82321.1; ALT_INIT.
DR EcoGene; EG11176; garR.
DR InterPro; IPR002204; 3hydroxisobut_dh.
DR InterPro; IPR006183; 6PGD.
DR InterPro; IPR006115; 6PGD_NAD.
DR Pfam; PFO3446; NAD binding 2; 1.
DR PRINTS; PR00076; 6PGDHRGNASE.
DR TIGRfams; TIGR01505; tartrosem_red; 1.
DR PROSITE; PS00895; 3_HYDROXYISOBUT_DH; 1.
DR Oxidoreductase; NAD; Complete proteome.
FT ACT_SITE 170 170 BY SIMILARITY.
SQ SEQUENCE 294 AA; 30427 MW; 17DA392C2253278C CRC64;

Query Match 12.8%; Score 367; DB 1; Length 294;
Best Local Similarity 29.4%; Pred. No. 3.8e-16;
Matches 83; Conservative 57; Mismatches 140; Indels 2; Gaps 1;

QY 269 KIGFLGLGLMGSGIVSNLLKMGHTVTWNRTAEKCDLFIQEGARLGTTPAEVVSTCDITF 328
DB 2 KVGFIGLGGKPKSKNLLKAGSLVVDNRNPAIADVIAGAEATASTAKIAEQCDVII 61
QY 329 ACVSDPKAAKDLVLPGSGVLQIRPKCYVDMSTVDADTVELAQVIVSGRGRFLAPVS 388
DB 62 TMLPNSPHVKEVALGENGIEGAKPGTVLIDMSSIAPLASREISEALKAGIDMLDAPVS 121
QY 389 GNOQLSNDGMVLVLAAGDRGLYEDSCSCFOAMGKTSFFLGEVGNAAKMMLIVNMVQSEF 448
DB 122 GGEFPAIDGTLSVMVGDKAIFDKYDLMKAMAGSVVHTGEIGAGNVTKLANQVIVALNI 181
QY 449 ATIAEGLTIAHVTGOSQQTLLDLINQGLASIFLDQKQCNILQGNFKPFPYLYKIQDLR 508
DB 182 AMSEALTATKAGVNDLVQAIRGLAGSLTLDKAKPMVMDNRNFKPGPRIDLHIKDLA 241
QY 509 LATALGDVNHPTPMAAAANEVYK--RAKALDQSDNDMSAVY 548
DB 242 NALDTSHGVAQLPLTAAVVENMMQALRADGLGTADHSALACY 283

RESULT 3

```

```

Y229 SYNY3
ID Q55702; STANDARD; PRT; 290 AA.
AC Q55702;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical oxidoreductase slr0229 (EC 1.1.1.-.-).
GN SLR0229.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
[1]
SEQUENCE FROM N.A.
MEDLINE=96127529; PubMed=8590279;
RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA Sugiura M., Tabata S.;
RA "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 64% to 92% of the genome.";
RL DNA Res. 2:153-166(1995).
CC -1- SIMILARITY: Belongs to the 3-hydroxyisobutyrate dehydrogenase
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D64000; BAA10233.1; -.
DR PIR; S76381; S76381.
DR InterPro; IPR002204; 3hydroxisobut_dh.
DR InterPro; IPR006183; 6PGD.
DR Pfam; PFO3446; NAD binding 2; 1.
DR PRINTS; PR00076; 6PGDHRGNASE.
DR PROSITE; PS00895; 3_HYDROXYISOBUT_DH; 1.
DR Oxidoreductase; NAD; Complete proteome.
FT ACT_SITE 175 175 BY SIMILARITY.
SQ SEQUENCE 290 AA; 29930 MW; BB167CD4FCF4A37A CRC64;

Query Match 11.8%; Score 338; DB 1; Length 290;
Best Local Similarity 29.6%; Pred. No. 2.5e-14;
Matches 85; Conservative 50; Mismatches 144; Indels 8; Gaps 3;

QY 269 KIGFLGLGLMGSGIVSNLLKMGHTVTWNRTAEKCDLFIQEGARLGTTPAEVVSTC 324
DB 5 KIAVFGLVGWSGPMQNLVKNQYQTVGNRTLRPS--VQEAAGKAGVKVVTIAVAANA 62
QY 325 DITFACVSDPKAAKDLVLPGSGVLQIRPKCYVDMSTVDADTVELAQVIVSGRGRFLE 384
DB 63 DITLTGCGDEKDVQQLILGSGGIAEYAKPQALIDCTIGTKTAAYELATNLKQLGLRFLD 122
QY 385 APVSGNQQLSNDGMVLVLAAGDRGLYEDSCSCFOAMGKTSFFLGEVGNAAKMMLIVNMVQ 444
DB 123 APTGSDVGAINTLTIWVGDISDFEALPVUKSIGEKIVHCGPSGSGQAVLNCQVLC 182
QY 445 GSFMTATIAEGLTIAHVTGOSQQTLLDLINQGLASIFLDQKQCNILQGNFKPFPYLYKIQ 504
DB 183 GIHAIAAEASAIQLSEQIGIAPLVIDTCGSGRAGSNALTNLAPKMGSEADPAPGMVKHLL 242
QY 505 KDLRLAIALGDVNHPTPMAAAANEVYKRAKALDQSDNDMSAVYRAY 551
DB 243 KDLRLVREA--AENGPLPGVTIAESLFTSVQLIGGEDQGSQAIRAY 287

RESULT 4
D3HI_ARATH
ID D3HI_ARATH STANDARD; PRT; 347 AA.
AC Q9SUC0; Q8LC25;
DT 16-OCT-2001 (Rel. 40, Created)

```



```

CC CC -!- CATALYTIC ACTIVITY: 3-hydroxy-2-methylpropanoate + NAD(+) = 2-
CC CC methyl-3-oxopropanoate + NADH.
CC CC -!- SUBUNIT: Homodimer (By similarity).
CC CC -!- SUBCELLULAR LOCATION: Mitochondrial (By similarity).
CC CC -!- SIMILARITY: Belongs to the 3-hydroxyisobutyrate dehydrogenase
CC CC family.
CC CC
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC
CC CC EMBL; BC003914; AA03914.1; -.
CC CC EMBL; AK078175; BAC37162.1; -.
CC CC PIR; PT0534; PT0534.
CC CC MGI; 1889802; 6430402H1ORik.
CC CC DR InterPro; IPR002204; 3hydroxisobut_dh.
CC CC DR InterPro; IPR006183; 6PGD.
CC CC DR Pfam; PF03446; NAD binding_2; 1.
CC CC DR PRINTS; PR00076; 6PGDHRGNASE.
CC CC DR PROSITE; PS00895; 3 HYDROXYISOBUT DH; 1.
CC CC KW Oxidoreductase; NAD; Mitochondrion; Transit peptide.
CC CC FT CHAIN 1 35 MITOCHONDRION (BY SIMILARITY).
CC CC FT NP BIND 36 335 3-HYDROXYISOBUTYRATE DEHYDROGENASE.
CC CC FT ACT_SITE 208 208 NAD (ADP PART) (POTENTIAL).
CC CC FT CONFLICT 2 2 BY SIMILARITY.
CC CC FT CONFLICT 2 2 A -> S (IN REF. 1).
CC CC SQ SEQUENCE 335 AA; 35440 MW; 5E9ECB03997DB110 CRC64;
CC CC
CC CC Query Match 10.4%; Score 297; DB 1; Length 335;
CC CC Best Local Similarity 26.1%; Pred. No. 1.1e-11;
CC CC Matches 81; Conservative 58; Mismatches 157; Indels 14; Gaps 4;
CC CC
CC CC QY 247 GSTSIQAADSTAVNGSITPTDKKIGFLGLMGSGIVSNLLKMGHTVTVMNRTAEKCDLF 306
CC CC Db 25 GSIAAVCSRSMA---SKTP---VGFGLGNMGNPMKMLKKGYPILLYDVPDVCKEF 77
CC CC
CC CC QY 307 IOEGARLGRTPAEVWSTCDITFACVSDPKAAKDLVLGSPVQLGIRPKCYVDMSTVDAD 366
CC CC Db 78 KEAGEQVASSPADVAEKADRIITMLPSSMNVAVSYSGANGILKKVKGSLIDSITDPS 137
CC CC
CC CC QY 367 IVTELAQVIVSGRFLFAPVSGNQQLSNDGMVILAAAGDRGLYEDCSCFOAMGKTSFF 426
CC CC Db 138 VSKELAKEVEKMGAVFMDAPVSGVGAARSGNLTFMVGGVEDEFAAAQELLECMSGNVVY 197
CC CC
CC CC QY 427 LGEVGNAAKMLVNVVQGSFMTATIEGLTLAHTVCSQOQLDILN--QGOLASIFLQ 484
CC CC Db 198 CGAVGTGQSAKICNNMLLAISMTGTAEAMNLGIRSLDPLKALIKLNMSSGRCSSTDY 257
CC CC
CC CC QY 485 KCQNILQG-----NFKPFDLYIKQDLRLAIALGDAVNHPTPMAAAANEVYKAKALQ 539
CC CC Db 258 PFGVMHGVPSNNYQGGFTLMAKDLGLAQDSATSTKTPILLGLSLAHQIYRMCKSGY 317
CC CC
CC CC QY 540 SNDMSAVYR 549
CC CC Db 318 SKKDFSSVFQ 327
CC CC
CC CC RESULT 7
CC CC D3HI_RAT
CC CC ID D3HI_RAT STANDARD; PRT; 335 AA.
CC CC AC P29266;
CC CC DT 01-DEC-1992 (Rel. 24, Created)
CC CC DT 15-OCT-2001 (Rel. 40, Last sequence update)
CC CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC CC DE 3-hydroxyisobutyrate dehydrogenase, mitochondrial precursor
CC CC DE (EC 1.1.1.31) (HIBADH).
CC CC GN HIBADH.
CC CC OS Rattus norvegicus (Rat).

```

```

OC CC -!- CATALYTIC ACTIVITY: 3-hydroxy-2-methylpropanoate + NAD(+) = 2-
OC CC methyl-3-oxopropanoate + NADH.
OC CC -!- SUBUNIT: Homodimer (By similarity).
OC CC -!- SUBCELLULAR LOCATION: Mitochondrial (By similarity).
OC CC -!- SIMILARITY: Belongs to the 3-hydroxyisobutyrate dehydrogenase
OC CC family.
OC CC
OC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
OC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
OC CC the European Bioinformatics Institute. There are no restrictions on its
OC CC use by non-profit institutions as long as its content is in no way
OC CC modified and this statement is not removed. Usage by and for commercial
OC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
OC CC or send an email to license@isb-sib.ch).
OC CC
OC CC EMBL; BC003914; AA03914.1; -.
OC CC EMBL; AK078175; BAC37162.1; -.
OC CC PIR; PT0534; PT0534.
OC CC MGI; 1889802; 6430402H1ORik.
OC CC RP InterPro; IPR002204; 3hydroxisobut_dh.
OC CC RP InterPro; IPR006183; 6PGD.
OC CC RP Pfam; PF03446; NAD binding_2; 1.
OC CC RP PRINTS; PR00076; 6PGDHRGNASE.
OC CC RP PROSITE; PS00895; 3 HYDROXYISOBUT DH; 1.
OC CC KW Oxidoreductase; NAD; Mitochondrion; Transit peptide.
OC CC FT CHAIN 1 35 MITOCHONDRION (BY SIMILARITY).
OC CC FT NP BIND 36 335 3-HYDROXYISOBUTYRATE DEHYDROGENASE.
OC CC FT ACT_SITE 208 208 NAD (ADP PART) (POTENTIAL).
OC CC FT CONFLICT 2 2 BY SIMILARITY.
OC CC FT CONFLICT 2 2 A -> S (IN REF. 1).
OC CC SQ SEQUENCE 335 AA; 35440 MW; 5E9ECB03997DB110 CRC64;
OC CC
OC CC Query Match 10.4%; Score 297; DB 1; Length 335;
OC CC Best Local Similarity 26.1%; Pred. No. 1.1e-11;
OC CC Matches 81; Conservative 58; Mismatches 157; Indels 14; Gaps 4;
OC CC
OC CC QY 247 GSTSIQAADSTAVNGSITPTDKKIGFLGLMGSGIVSNLLKMGHTVTVMNRTAEKCDLF 306
OC CC Db 25 GSIAAVCSRSMA---SKTP---VGFGLGNMGNPMKMLKKGYPILLYDVPDVCKEF 77
OC CC
OC CC QY 307 IOEGARLGRTPAEVWSTCDITFACVSDPKAAKDLVLGSPVQLGIRPKCYVDMSTVDAD 366
OC CC Db 78 KEAGEQVASSPADVAEKADRIITMLPSSMNVAVSYSGANGILKKVKGSLIDSITDPS 137
OC CC
OC CC QY 367 IVTELAQVIVSGRFLFAPVSGNQQLSNDGMVILAAAGDRGLYEDCSCFOAMGKTSFF 426
OC CC Db 138 VSKELAKEVEKMGAVFMDAPVSGVGAARSGNLTFMVGGVEDEFAAAQELLECMSGNVVY 197
OC CC
OC CC QY 427 LGEVGNAAKMLVNVVQGSFMTATIEGLTLAHTVCSQOQLDILN--QGOLASIFLQ 484
OC CC Db 198 CGAVGTGQSAKICNNMLLAISMTGTAEAMNLGIRSLDPLKALIKLNMSSGRCSSTDY 257
OC CC
OC CC QY 485 KCQNILQG-----NFKPFDLYIKQDLRLAIALGDAVNHPTPMAAAANEVYKAKALQ 539
OC CC Db 258 PFGVMHGVPSNNYQGGFTLMAKDLGLAQDSATSTKTPILLGLSLAHQIYRMCKSGY 317
OC CC
OC CC QY 540 SNDMSAVYR 549
OC CC Db 318 SKKDFSSVFQ 327
OC CC
OC CC RESULT 7
OC CC D3HI_RAT
OC CC ID D3HI_RAT STANDARD; PRT; 335 AA.
OC CC AC P29266;
OC CC DT 01-DEC-1992 (Rel. 24, Created)
OC CC DT 15-OCT-2001 (Rel. 40, Last sequence update)
OC CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
OC CC DE 3-hydroxyisobutyrate dehydrogenase, mitochondrial precursor
OC CC DE (EC 1.1.1.31) (HIBADH).
OC CC GN HIBADH.
OC CC OS Rattus norvegicus (Rat).

```

```

OC CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC CC NCBI_TaxID=10116;
OC CC [1]
OC CC SEQUENCE FROM N.A.
OC CC STRAIN=Sprague-Dawley; TISSUE=Liver;
OC CC MEDLINE=89174651; PubMed=2647728;
OC CC RA Rougaff P.M., Zhang B., Kuntz M.J., Harris R.A., Crabb D.W.;
OC CC "Cloning and sequence analysis of a cDNA for 3-hydroxyisobutyrate
OC CC dehydrogenase. Evidence for its evolutionary relationship to other
OC CC pyridine nucleotide-dependent dehydrogenases.";
OC CC RL J. Biol. Chem. 264:5899-5903(1989).
OC CC [2]
OC CC MUTAGENESIS.
OC CC RP MEDLINE=96335606; PubMed=8766712;
OC CC RA Hawes J.W., Harper E.T., Crabb D.W., Harris R.A.;
OC CC "Structural and mechanistic similarities of 6-phosphogluconate and 3-
OC CC hydroxyisobutyrate dehydrogenases reveal a new enzyme family, the 3-
OC CC hydroxyacid dehydrogenases.";
OC CC RL FEBS Lett. 389:263-267(1996).
OC CC [3]
OC CC CONCEPTUAL TRANSLATION.
OC CC RP Bairoch A.;
OC CC UNPUBLISHED OBSERVATIONS (JAN-2001).
OC CC RL -!- CATALYTIC ACTIVITY: 3-hydroxy-2-methylpropanoate + NAD(+) = 2-
OC CC methyl-3-oxopropanoate + NADH.
OC CC -!- SUBUNIT: Homodimer.
OC CC -!- SUBCELLULAR LOCATION: Mitochondrial.
OC CC -!- TISSUE SPECIFICITY: HIGHER LEVEL IN KIDNEY, LIVER, AND HEART THAN
OC CC IN MUSCLE.
OC CC -!- SIMILARITY: Belongs to the 3-hydroxyisobutyrate dehydrogenase
OC CC family.
OC CC -!- CAUTION: Ref.1 sequence differs from that shown due to a
OC CC frameshift in position 10.
OC CC
OC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
OC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
OC CC the European Bioinformatics Institute. There are no restrictions on its
OC CC use by non-profit institutions as long as its content is in no way
OC CC modified and this statement is not removed. Usage by and for commercial
OC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
OC CC or send an email to license@isb-sib.ch).
OC CC
OC CC EMBL; J04628; AAA50312.1; ALT_FRAME.
OC CC PIR; A32867; A32867.
OC CC DR InterPro; IPR002204; 3hydroxisobut_dh.
OC CC DR InterPro; IPR006183; 6PGD.
OC CC DR InterPro; IPR006115; 6PGD_NAD.
OC CC DR Pfam; PF03446; NAD binding_2; 1.
OC CC DR PRINTS; PR00076; 6PGDHRGNASE.
OC CC DR PROSITE; PS00895; 3 HYDROXYISOBUT DH; 1.
OC CC KW Oxidoreductase; NAD; Mitochondrion; Transit peptide.
OC CC FT CHAIN 1 35 MITOCHONDRION (BY SIMILARITY).
OC CC FT NP BIND 36 335 3-HYDROXYISOBUTYRATE DEHYDROGENASE.
OC CC FT ACT_SITE 208 208 NAD (ADP PART) (POTENTIAL).
OC CC FT MUTAGEN 68 68 D->R: DECREASE OF ACTIVITY WITH NAD,
OC CC INCREASE OF ACTIVITY WITH NADP.
OC CC FT MUTAGEN 208 208 K->A,H,N,R: COMPLETE LOSS OF ACTIVITY.
OC CC FT MUTAGEN 212 212 N->Q: DECREASE IN ACTIVITY.
OC CC SQ SEQUENCE 335 AA; 35302 MW; D266A7838500295A CRC64;
OC CC
OC CC Query Match 10.1%; Score 290; DB 1; Length 335;
OC CC Best Local Similarity 25.5%; Pred. No. 3e-11;
OC CC Matches 79; Conservative 60; Mismatches 157; Indels 14; Gaps 4;
OC CC
OC CC QY 247 GSTSIQAADSTAVNGSITPTDKKIGFLGLMGSGIVSNLLKMGHTVTVMNRTAEKCDLF 306
OC CC Db 25 GSIAAVCSRSMA---SKTP---VGFGLGNMGNPMKMLKKGYPILLYDVPDVCKEF 77
OC CC
OC CC QY 307 IOEGARLGRTPAEVWSTCDITFACVSDPKAAKDLVLGSPVQLGIRPKCYVDMSTVDAD 366
OC CC Db 78 KEAGEQVASSPADVAEKADRIITMLPSSMNVAVSYSGANGILKKVKGSLIDSITDPS 137

```

```

QY 367 TVTELAQVIVSGRFLAPVSGNQSLNDGMLVILAAGDRGLYEDCSCFOAMGKTSPF 426
D 138 VSKELAKEVERKMGAVFMDAPVSGVGAARSGNLTFTMVGGVENEFAAQELLGCMGSLVLY 197
QY 427 LGEVGNAAKMMMLVNMVQGSFNATTAAGTLTAHVTGQSQOQLDILN--QGQLASIFLDQ 484
D 198 CAVGSGGASAKTCNNMLLAISMGITFAEAMNGLGIRSLDPKLLAKLINMSSGRCSSTYN 257
QY 485 KCQNILQ-----NFKPDPFYLYKIQKDLRLATLADGVNHPPTMAAAANEVVKRAKALDQ 539
D 258 PVPGVMDGVPSNNYQGGFTTLMADKDLGLAQDSATSTKTPILLGSAVHQIYRMCMCKGY 317
QY 540 SDNDMSAVYR 549
D 318 SKKDFSSVFQ 327

RESULT 8
YIHU_SALTY STANDARD; PRT; 298 AA.
AC Q9L7S0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical oxidoreductase yihU (EC 1.1.-.-).
GN YIHU OR STM4023.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Lt2;
RA Krogan N.J., Zhang R., Neuhaud J., Kelln R.A.;
RT "Utilization of dihydroxotate as sole pyrimidine source by Salmonella
  typhimurium.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Lt2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
  Lt2.";
RL Nature 413:852-856(2001).
CC -1- SIMILARITY: Belongs to the 3-hydroxyisobutyrate dehydrogenase
  family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC ENBL; AF220438; AAF27921.1; -;
CC EMBL; AE008887; AAL22862.1; -;
CC StvGene; SG77777; YihU.
CC InterPro; IPR002204; 3hydroxisobut_dh.
CC InterPro; IPR006115; 6PGD_NAD.
CC InterPro; IPR000205; NAD_BS.
CC Pfam; PF03446; NAD_binding_2; 1.
CC PROSITE; PS00895; 3 HYDROXYISOBUT DH; 1.
CC Hypothetical protein; Oxidoreductase; NAD; Complete proteome.
CC ACT SITE 171 171 BY SIMILARITY.
CC SEQUENCE 298 AA; 31280 MW; 7ED9FA174249240C CRC64;

```

```

Query Match 10.1%; Score 289.5; DB 1; Length 298;
Best Local Similarity 25.4%; Pred. No. 2.8e-11;
Matches 72; Conservative 59; Mismatches 143; Indels 9; Gaps 2;

QY 270 IGFIGLGLMGSGIVSNLLKMGHTVTVMNRTEAKCDLFTOEAGRLGRTPAEVVTSTCDITFA 329
D 4 IAFIGLGMGSPMASNLLKQGHLSFVDPNPAVQLVNDKGAQPASSPAQAATIGAEEFVIT 63
QY 330 CVSPDKAAKDLVLGSPGVLOGRPKCVDMSTVDADVTVELAQVIVSRGGRFLRAPVSG 389
D 64 MLPNGDLVRSVLFESGQGVCELTSLREALVIDMSTTHPLQTDNLADIADMQSGKFSMDVPFGR 123
QY 390 NQOALSNDGMLVILAAGDRGLYEDCSCFOAMGKTSTFELGVEVGNAAKMMMLVNMVQGSFWA 449
D 124 TSDNATITGTLILLAGGTAEQVERATPVLMANGNELVNTGGPGMGIRVRLINNYMSIALNA 183
QY 450 TIAEGLTLAHVTGOSQOQLDILN-----QGQLASIFLDQKCNILQGNFKPDPFYLYKVIQ 504
D 184 LSAEAAVLCEALGLSFDVALKVMGSTAAGKGHFTTWPNK----VMKGDLSFAFMIDLAH 239
QY 505 KDLRLATLADGVNHPPTMAAAANEVVKRAKALDQSDNDMSAV 547
D 240 KDLGIALDVANQLHVPMPGLGAASREVYNLARAAGRGREDWSAI 282

RESULT 9
YIHU_ECOLI STANDARD; PRT; 298 AA.
AC P32142;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical oxidoreductase yihU (EC 1.1.-.-).
GN YIHU OR B3882 OR SF3954 OR S3792.
OS Escherichia coli, and
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 623;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12 / MG1655;
RX MEDLINE=93347969; PubMed=8346018;
RA Plunkett G. III, Burland V.D., Daniels D.L., Blattner F.R.;
RT "Analysis of the Escherichia coli genome. III. DNA sequence of the
  region from 87.2 to 89.2 minutes.";
RL Nucleic Acids Res. 21:3391-3398(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao Y., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
  through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=S.flexneri; STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
  flexneri serotype 2a strain 2457T.";
RL Infect. Immun. 71:2775-2786(2003).
CC -1- SIMILARITY: Belongs to the 3-hydroxyisobutyrate dehydrogenase
  family.
CC

```

```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L19201; AAB03015.1; -.
DR EMBL; AE000464; AAD13444.1; -.
DR EMBL; AE015402; AAN45389.1; -.
DR EMBL; AE016990; AAP18811.1; -.
DR PIR; S40826; S40826.
DR EcoGene; EG11847; YihU.
DR InterPro; IPR002204; 3hydroxisobut_dh.
DR InterPro; IPR006115; 6PGD_NAD.
DR Pfam; PF03446; NAD_binding_2; 1.
DR PROSITE; PS00895; 3_HYDROXYISOBUT_DH; 1.
KW Hypothetical protein; Oxidoreductase; NAD; Complete proteome.
FT ACT_SITE 171 171 BY SIMILARITY.
SQ SEQUENCE 298 AA; 31158 MW; 74FEC8C09FA7881C CRC64;
Query Match 10.0%; Score 287.5; DB 1; Length 298;
Best Local Similarity 26.4%; Pred. No. 3.7e-11;
Matches 74; Conservative 54; Mismatches 149; Indels 3; Gaps 2;
QY 270 IGFLGLMGSGIVSNLLKMGHTVTVWNTAKCDLFTQEGARLGRTPAEVYSTCDITFA 329
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4 IAFIGLQMGSPMASNLLQQGHLRFVDVNAEVRHLVDKGATPAANPAQAQAKDAEFIT 63
QY 330 CYSVDPAKADLVLPQGVVQVLTGTRPGKCVYDMSTVDADTVTELAQVTVSGRFLBAPVSG 389
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
64 MLPNGLVRNVLFGNGVCEGLSTDALVIDMSTIIHPLQTDKLADMQAGFPMVPGR 123
QY 390 NQQLSNDGMLVLAAGDRGLYEDSCSCFOAMGKTSEFFLGEVGNAAKMMILVNVQGSFWA 449
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
124 TSANAITGTLILLAGTAGBQVERATPILMANGSELINAGPGGIRVKVLNNYMSIALNA 183
QY 450 TIAEGTLIAHTVGTGSGQTLIDILNOGQLA--SIFLDQKCNILQGNFKDFVLYKIQKDL 507
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
184 LSAAEAVLCEALNLPFDVAVKVM--GTAAGKGHTTSWPNKVLSGDLSPAFMDLAHKDL 242
QY 508 RLALGADVAHNPETMAAANEVYKRAKALDOSDNDMSAV 547
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
243 GIALDVANQLHVMPGLGAASREVYSQARAAGRQDWSAI 282
RESULT 10
D3H1_HUMAN STANDARD; PRT; 336 AA.
AC P31937; Q9UDN3;
DT 01-JUL-1993 (Rel. 26, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 2-hydroxyisobutyrate dehydrogenase, mitochondrial precursor
DE (EC 1.1.1.31) (HIBADH).
GN HIBADH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.M., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmitt J., Myers R.M.,
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Marra M.A.;
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences".;
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE OF 32-336 FROM N.A.
RA Du F., Wohlmann P., Holmes A.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 37-47.
RX TISSUE=Liver;
RC MEDLINE=94147969; PubMed=8313870;
RA Hughes G.J., Frutiger S., Paquet N., Pasquali C., Sanchez J.-C.,
RA Tissot J.-D., Bairoch A., Appel R.D., Hochstrasser D.F.;
RT "Human liver protein map: update 1993.";
CC Electrophoresis 14:1216-1222(1993).
CC -!- CATALYTIC ACTIVITY: 3-hydroxy-2-methylpropanoate + NAD(+) = 2-
CC methyl-3-oxopropanoate + NADH.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Mitochondrial.
CC -!- SIMILARITY: Belongs to the 3-hydroxyisobutyrate dehydrogenase
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; BC032324; AAH32324.1; -.
DR EMBL; AC007130; -- NOT ANNOTATED_CDS.
DR SWISS-2DPAGE; P31937; HUMAN.
DR Sienna-2DPAGE; P31937; -.
DR Genew; HGNC:4907; HIBADH.
DR GK; P31937; -.
DR MIM; 236795; -.
DR GO; GO:0005739; C-mitochondrion; NAS.
DR GO; GO:0008442; F:3-hydroxyisobutyrate dehydrogenase activity; NAS.
DR GO; GO:0006573; P:valine metabolism; NAS.
DR InterPro; IPR002204; 3hydroxisobut_dh.
DR InterPro; IPR006115; 6PGD.
DR InterPro; IPR006115; 6PGD_NAD.
DR Pfam; PF03446; NAD_binding_2; 1.
DR PRINTS; PR00076; 6PGDHNGASE.
DR PROSITE; PS00895; 3_HYDROXYISOBUT_DH; 1.
KW Oxidoreductase; NAD; Mitochondrion; Transit peptide.
FT TRANSIT 1 36 MITOCHONDRION.
FT CHAIN 37 336 3-HYDROXYISOBUTYRATE DEHYDROGENASE.
FT NP_BIND 40 68 NAD (ADP PART) (POTENTIAL).
FT ACT_SITE 209 209 BY SIMILARITY.
SQ SEQUENCE 336 AA; 35329 MW; DA3128774A91AF48 CRC64;
Query Match 9.8%; Score 281; DB 1; Length 336;
Best Local Similarity 26.4%; Pred. No. 1.1e-10;
Matches 81; Conservative 53; Mismatches 161; Indels 12; Gaps 4;
QY 250 SIQAADSTAVNGSITPTDKKIGFLGLMGSGIVSNLLKMGHTVTVWNTAKCDLFIQE 309
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
27 SFAAVCSRSV-ASKTP---VGFILGNMGPNMAKMLKMGYFLIIVDFPDACKEFQDA 81
QY 310 GARLGRTPAEVYSTCDITFACVSDPKAAKDLVLPGSVLQIRPGKCYVDMSTVDADTVT 369
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
82 GEQVSSPDAEAKADRIITMLPTSINAIEYSGANGILKVKKGLLIDSTIDPAVSK 141
```

Qy	370	ELAQVIVSRGGRFLEAPVSGNQQLSDNGMLVILAAAGDRGLYEDCSCSFQAMGKTSFFFLGE	420
Db	142	ELAKEVERKMGAVFMDAPVSGVGGVGAARSGNLTFTMVGSGVEDEFAAAQELLGCMGNSVYVCGA	201
Qy	430	VGNAAKMWLVINVMVQGSFWATIAEGLTLAHVTGQSQQTLLDLN--QGOLASIFLDQKQC	487
Db	202	VGTQAAKICNNMLLAISMIGTAAEAMNIGIRLGLDPLKLLAKTLNNSSRCWSSDTYNPVP	261
Qy	488	NILOG-----NFKPDFLYLKQKOLRLAIALGDVNVHPTPEMAAANENVVKRAKALDQSDN	542
Db	262	GVMDGVFSANNYQGGFTTLLAKDLGLAQDSNTSKSPILGSLAHQIYRMCAKGYSKK	321
Qy	543	DMSAVYR	549
Db	322	DFSSVFQ	328

RESULT 11

D3HI_DROME

ID

D3HI_DROME

STANDARD;

PRT;

324 AA.

QV8M5;

Q9V8M6;

AC

16-OCT-2001

(Rel. 40, Created)

DT

16-OCT-2001

(Rel. 40, Last sequence update)

DT

10-OCT-2003

(Rel. 42, Last annotation update)

DT

Probable 3-hydroxyisobutyrate dehydrogenase, mitochondrial precursor

DE

(EC 1.1.1.31) (HIBADH).

DE

CGI5093.

GN

OS

Drosophila melanogaster (Fruit fly).

OS

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC

Ephydroidea; Drosophilidae; Drosophila.

OX

NCBI_TaxID=7277;

RP

[1]

RP

SEQUENCE FROM N.A.

RC

STRAIN=Berkely;

RC

MEDLINE=20136006; PubMed=10731132;

RA

Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA

Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA

George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA

Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA

Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

RA

Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA

Abri J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA

Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA

Beeson K.Y., Benos P.V., Berman B.P., Brokstein P., Brotter P.,

RA

Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,

RA

Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA

de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA

Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA

Durbin K.J., Evangelista C.C., Ferraz C., Davenport L.B., Davies P.,

RA

Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA

Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA

Hartis N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,

RA

Hosdin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA

Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA

Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA

Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,

RA

Liu X., Mantei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA

Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA

Mout S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA

Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,

RA

Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA

Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA

Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA

Spier E., Spradling A.C., Scapleton M., Strong R., Sun E.,

RA

Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA

Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,

RA

Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA

Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA

Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA

Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

```

KW Oxidoreductase; NAD; Complete proteome.
FT NP_BIND 3 31 NAD (BY SIMILARITY).
FT ACT_SITE 171 BY SIMILARITY.
SQ SEQUENCE 298 AA; 30486 MW; 0C4D7B5A7C870730 CRC64;

Query Match 9.1%; Score 261.5; DB 1; Length 298;
Best Local Similarity 25.8%; Pred. No. 1.6e-09;
Matches 76; Conservative 51; Mismatches 151; Indels 17; Gaps 4;

QY 270 IGFGLGLMGSGIVSNLLKMGHTVTVNRTAEKCDLFTQEGARLGRTPAEVVSTCDITFA 329
DB 4 IAFGLGNGMGPMANLLKAGHRVNFDPKAVLGLVEQAGQADGADSLQCEGAEEVVIS 63
QY 330 CVSPQKAAKDLVLGPGSVLQIRPKCYVDMSTVDADITVTELAQVIVSRGGRFLBAPVSG 389
DB 64 MLPAGQHVESLYLGGDGLLARVAGKPLLDICSTIAPTARKVAEAAAKGLTLLDAPVSG 123
QY 390 NOQLSNDGMVLAAAGRLGYEDCSSCFQAMGKTSFFLGEVGNAAKMLLVNMGVSFMA 449
DB 124 GVGGARAGTSLFVGGPAGFARAPVLENMGRNIFHAGDHGAGQVAKICNNMLLGILMA 183
QY 450 TIAEGLTLAHVTGQSQQTLLDILNQ---GQLA-----SIFLDQKQNILQGNFKPDF 498
DB 184 GTAEALALGVKNGLDPAVLSEVMKQSSGNGWALNLYNPWGPQAPASNGVAGGFQ--- 240
QY 499 YLKTIQKDLRLAIALGDVNHPTMAAAANEVYKRAKALDQSDN--DMSAVTRAY 551
DB 241 -VRLNKKDLGLALANAQAQVASTPLGLARNLFLSHAQADAHEGLDFDSIQKLY 294

RESULT 13
DH31 CAEEL STANDARD; PRT; 299 AA.
ID DH31 CAEEL STANDARD; PRT; 299 AA.
AC Q9XTT0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable 3-hydroxyisobutyrate dehydrogenase, mitochondrial
DE (EC 1.1.1.31) (HIBADH).
GN B0250.5
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OC NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Lloyd C., White S.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC -|- CATALYTIC ACTIVITY: 3-hydroxy-2-methylpropanoate + NAD(+) = 2-
CC methyl-3-oxopropanoate + NADH.
CC -|- SUBCELLULAR LOCATION: Mitochondrial (By similarity).
CC -|- SIMILARITY: Belongs to the 3-hydroxyisobutyrate dehydrogenase
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z81453; CAB03798.1; -.
CC DR EMBL; AL031630; CAB03798.1; JOINED.
CC DR EMBL; AL031630; CAB03798.1; -.
CC DR EMBL; Z81453; CAB03798.1; JOINED.
CC DR PIR; T18682; T18682.
CC DR WormPep; B0250.5; CE18481.
CC DR InterPro; IPR002204; 3hydroxisobut_dh.
CC DR PROSITE; PS00895; 3_HYDROXYISOBUT_DH; 1.
CC KW Hypothetical protein; Oxidoreductase; NAD; Mitochondrion.
CC FT NP_BIND 2 30 NAD (ADP PART) (POTENTIAL).

```

RESULT 12

```

MMSB_PSEAE STANDARD; PRT; 298 AA.
ID MMSB_PSEAE STANDARD; PRT; 298 AA.
AC P28811;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 3-hydroxyisobutyrate dehydrogenase (EC 1.1.1.31) (HIBADH).
GN MMSB OR PA3569.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OC NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=1339433;
RA Steele M.I., Lorenz D., Hatter K., Park A., Sokatch J.R.;
RA "Characterization of the mmsAB operon of Pseudomonas aeruginosa PAO
RA encoding methylmalonate-semialdehyde dehydrogenase and 3-
RA hydroxyisobutyrate dehydrogenase."
RA J. Biol. Chem. 267:13585-13592(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Harbig K., Lim R.W.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RA "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RA opportunistic pathogen."
RA Nature 406:959-964(2000).
RL Nature 406:959-964(2000).
CC -|- CATALYTIC ACTIVITY: 3-hydroxy-2-methylpropanoate + NAD(+) = 2-
CC methyl-3-oxopropanoate + NADH.
CC -|- PATHWAY: Distal valine metabolic pathway.
CC -|- INDUCTION: By valine.
CC -|- SIMILARITY: Belongs to the 3-hydroxyisobutyrate dehydrogenase
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M84911; AAA25892.1; -.
CC DR EMBL; A8004778; AAG06957.1; -.
CC DR PIR; C42902; C42902.
CC DR InterPro; IPR002204; 3hydroxisobut_dh.
CC DR InterPro; IPR006115; 6PGD_NAD.
CC DR Pfam; PF03446; NAD binding 2; 1.
CC DR PROSITE; PS00895; 3_HYDROXYISOBUT_DH; 1.

```



```

FT ACT_SITE 172 172 PROBABLE.
SQ SEQUENCE 299 AA; 31217 MW; F30B423A597CABF0 CRC64;

Query Match
Best Local Similarity 9.1%; Score 259.5; DB 1; Length 299;
Matches 77; Conservative 48; Mismatches 143; Indels 31; Gaps 6;

QY 271 GFLGLGMSGIVSNLLKMGHTVTVNRTAECDFIOEGARLGRTPAEVSTCDITFAC 330
DB 5 GFIGLGNMGHMAKNLKNKKLIVDVNKAUVQEFKAEGCEVAARHPADIAAASKEIITV 64

QY 331 VSDPKAAKDLVLPSPGVLOGIRPGKCYVDMSTVDADTVELAQIVSRGGRLEAPVSGN 390
DB 65 LPSPSHVKAUVQGEAGIFKIQPGTCLMDSSTIDQIVSLEVAQAAALLKAEYDAPISGG 124

QY 391 ---QQLSNDGMLVILAAGDRLGYEDCSFCFQAMGKTSFFLGEVGNAAKMWLIIVNVOGS 446
DB 125 VTGAQAT---LTPVWGAGNDATFKRAEAVLSLMGKNVINLGAVGNGTAAKICNNMLIGI 181

QY 447 FMATIAEGLTIAHVTGSGQQTLDLIN-----QGOLASIFLDQKCNILQ 491
DB 182 QMVAVAETMNLGISMGLDAKALAGIVNTSSGRWSSDTYNPVPVGIENI---PSCRGYAG 238

QY 492 GNEKPDFVLKVIQDLRLAIALGDVNHPTMAAANAEVYK-RAKALDQSDNDMSAVTR 549
DB 239 G-----FGTLLMAKDLSLAQNASTNTQPTPMGSLAHQIYRILARDPOYQAKDFGVVYQ 292

RESULT 14
YGBJ ECOLI
ID YGBJ ECOLI STANDARD; PRT; 302 AA.
AC Q6888;
DT 15-JUL-1998 (Rel. 36, Created)
DE 15-JUL-1998 (Rel. 36, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical oxidoreductase ygbJ (EC 1.1.-.-).
GN YGBJ OR B2736.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RA "The complete genome sequence of Escherichia coli K-12.";
RT Science 277:1453-1474(1997).
RL -!- SIMILARITY: Belongs to the 3-hydroxyisobutyrate dehydrogenase
CC family.
CC
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC ENBL; U29579; AAA69246.1; --
CC DR ENBL; AE000357; AAC75778.1; --
CC DR PIR; D65054; D65054.
CC DR EcoGene; EG13104; ygbJ.
CC DR InterPro; IPR002204; 3hydroxisobut dh.
CC DR PROSITE; PS00895; 3 HYDROXYISOBUT DH; 1.
CC KW Hypothetical protein; Oxidoreductase; NAD; Complete proteome.
FT ACT_SITE 178 BY SIMILARITY
SQ SEQUENCE 302 AA; 30751 MW; 09E94256EDDFC26C CRC64;

Query Match
Best Local Similarity 22.9%; Pred. No. 4.1e-08;
Matches 67; Conservative 64; Mismatches 155; Indels 6; Gaps 4;

QY 264 PPTDKKIGFLGLGMSGIVSNLLKMGHTVTVN--NRTAECDFIOEGA-RLGRTPAEV 320
DB 3 TGSEPHVGVIGLGMGMGAALSYVRAG--LSTWGADLNSNACATLKEAGACGVSDNAATF 60

QY 321 VSTCDITFACVSDPKAAKDLVLPSPGVLOGIRPGKCYVDMSTVDADTVELAQIVSRGG 380
DB 61 AEKLDALLVLVNAQAQVQVLFGETGVAQHLKPGTAVMVSSITAGADAQEIATALAGFDL 120

QY 381 RFLEAPVSGNQQLSNDGMLVILAAGDRLGYEDCSFCFQAMGKTSFFLG-EVGNAAKMWLI 439
DB 121 EMLDAPVSGVKAANGEMTWVWSSDIAFERLAPVLEAVACKVYRIGAEPLGSGTVKII 180

QY 440 VNVQGSFMATIAEGLTIAHVTGSGQQTLDLINQGOLASIFLDQKCNILQGNFKPDFY 499
DB 181 HQLLAGVHIAAGAEAMALAAARAGIPLDVNYDVVTNNAAGSNWVFNMRHVVDGYTPHSA 240

QY 500 LKVIQDLRLAIALGDVNHPTMAAANAEVYKRAKALDQSDNDMSAVTRAY 551
DB 241 VDFVVKDLGLVADTAKALHFFPLPLASTALNMFTSASNAGYKGEDDSAVIKIF 292

RESULT 15
MMSB MYCTU
ID MMSB MYCTU STANDARD; PRT; 294 AA.
AC O53814;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Probable 3-hydroxyisobutyrate dehydrogenase (EC 1.1.1.31) (HIIBADH).
GN MMSB OR RV0751C OR MT0775 OR MT041.25C OR MB0773C.
OS Mycobacterium tuberculosis, and
OS Mycobacterium tuberculosis;
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773, 1765;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=M.bovis; STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972;
RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;

```

```

RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
CC -!- CATALYTIC ACTIVITY: 3-hydroxy-2-methylpropanoate + NAD(+) = 2-
CC methyl-3-oxopropanoate + NADH.
CC -!- PATHWAY: Distal valine metabolic pathway.
CC -!- SIMILARITY: Belongs to the 3-hydroxyisobutyrate dehydrogenase
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AL021958; CAAL7518.1; -
CC DR EMBL; AE006969; AAK45016.1; -
CC DR EMBL; BX248336; CAD93635.1; -
CC DR FIC; B70825; B70825.
CC TIGR; MT0775; -.
CC DR TubercuList; RV0751c; -.
CC DR InterPro; IPR002204; 3hydroxisobut_dh.
CC DR PROSITE; PS00895; 3-HYDROXYISOBUT_DH; 1.
CC KW Oxidoreductase; NAD; Complete proteome.
CC FT NPT_BIND 3 31 NAD (ADP PART) (POTENTIAL).
CC FT ACT_SITE 168 168 BY SIMILARITY.
CC SQ SEQUENCE 294 AA; 29679 MW; 68DEF129A8BA3E68 CRC64;
CC
CC Query Match 8.2%; Score 234.5; DB 1; Length 294;
CC Best Local Similarity 26.0%; Pred. No. 7.6e-08;
CC Matches 74; Conservative 39; Mismatches 161; Indels 11; Gaps 3;
CC
CC QY 270 IQFLGLGMLGSGIVNLLKMGHTVTVNRTAKCDLFIQEGARLGRTPAEVSTCDITFA 329
CC ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
CC Db 4 IAFGLGNGGAPMSANLVGAGHVRGDFDPAFTAAGAAAHGVAFPSAPEVAEADVVIT 63
CC
CC QY 330 CVSDPKAAKDLVLPGSGVLQGRPGKYVDMSTVDADTVELAQVIVSGRFLAPVSG 389
CC : : : : : : : : : : : : : : : : : : : : : : : : : :
CC Db 64 MLPTGTVVERCY----TDVLAARAPATLFDISSTISVTDAREVHALAESGMLQDAPVSG 120
CC
CC QY 390 NOOLNSNDGMLVILAAAGDRGLYEDSCSCFQAMGKTSFFLGEVGNAAKMLIVNVQGSFMA 449
CC : : : : : : : : : : : : : : : : : : : : : : : : : :
CC Db 121 GYKGAATAATLAFVWGDESTLRARPVLPEPMAGKITHCGAAGQAQAKVNNVLAQQI 180
CC
CC QY 450 TIAEGHTLAHVTCQSQOQLDILNQQLASIFLDQK-----QNLQGNFKPDFLYKY 502
CC : : : : : : : : : : : : : : : : : : : : : : : : : :
CC Db 181 ATAAEFVLAELKLSAQSLFDVITGATGNCWAVHTNCPVPGPVPTSPANNDFKPGFSTAL 240
CC
CC QY 503 IQXDLBLATLGDVAVNHPTMAAAANVEYKRAKALDOSNDMSAV 547
CC : : : : : : : : : : : : : : : : : : : : : : : : : :
CC Db 241 MNKDLGLMDAVAATGATAPLGSHAADIYAKF--AADHADLDLSAV 284
CC
CC RESULT 16
CC YFJR_BACSU STANDARD; PRT; 261 AA.
CC AC O34969;
CC DT 15-JUL-1998 (Rel. 36, Created)
CC DT 15-JUL-1998 (Rel. 36, Last sequence update)
CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
CC DE Hypothetical oxidoreductase yfjr (EC 1.1.-.-).
CC GN YFJR OR BSU07990.
CC OS Bacillus subtilis.
CC OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
CC NCBI_TaxID=1423;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=AC327;
CC RX MEDLINE=971124190; PubMed=8969503;
CC RA Yamamoto H., Uchiyama S., Sekiguchi J.;
CC RT "Cloning and sequencing of a 40.6 kb segment in the 73 degrees-76
CC degrees region of the Bacillus subtilis chromosome containing genes

```

```

for trehalose metabolism and acetoin utilization.";
RL Microbiology 142:3057-3065(1996).
CC [2]
CC SEQUENCE FROM N.A.
CC STRAIN=168;
CC MEDLINE=98044033; PubMed=9384377;
CC RA Kunst P., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
CC Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
CC Brouillet S., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
CC Choi S.K., Codani J.J., Comerton I.F., Cummings N.J., Daniel R.A.,
CC Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
CC Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
CC Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
CC Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
CC Guisepi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
CC Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
CC Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
CC Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
CC Kurita K., Lapidus A., Lardinio S., Lauber J., Lazarevic C.,
CC Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
CC Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
CC Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
CC Parro V., Pohl I.M., Portetelle D., Porwollik S., Prescott A.M.,
CC Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
CC Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
CC Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
CC Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
CC Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
CC Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
CC Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
CC Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
CC Winters P., Wipat K., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
CC Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
CC "The complete genome sequence of the Gram-positive bacterium Bacillus
CC subtilis.";
CC RL Nature 390:249-256(1997).
CC -!- SIMILARITY: Belongs to the 3-hydroxyisobutyrate dehydrogenase
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D83967; BAA23388.1; -
CC DR EMBL; D78509; BAA24303.1; -
CC DR EMBL; Z99108; CAB12628.1; -
CC DR PIR; A69807; A69807.
CC DR Subtilist; BG12914; yfjr.
CC DR InterPro; IPR002204; 3hydroxisobut_dh.
CC DR PROSITE; PS00895; 3-HYDROXYISOBUT_DH; FALSE_NEG.
CC KW Hypothetical protein; Oxidoreductase; NAD; Complete proteome.
CC FT ACT_SITE 146 146 BY SIMILARITY.
CC SQ SEQUENCE 261 AA; 27866 MW; 6C9A8CAC8C71CA66 CRC64;
CC
CC Query Match 7.8%; Score 224.5; DB 1; Length 261;
CC Best Local Similarity 24.9%; Pred. No. 2.8e-07;
CC Matches 66; Conservative 51; Mismatches 133; Indels 15; Gaps 5;
CC
CC QY 293 VTVNRTAKCDLFIQEGARLGRTPAEVSTCDITFAVCSDPKAAKDLVLPGSGVLQGR 352
CC ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
CC Db 1 MTVNRTKQKTDLVTEGAQAADTPRLAKSADIVITMLADDSVSTVTFEGDGLLEGLA 60
CC
CC QY 353 PGKYVDMSTVDADTVELAQVIVSGRFLAPVSGNQQLNSNDGMLVILAAAGDRGLYED 412
CC : : : : : : : : : : : : : : : : : : : : : : : : : :
CC Db 61 ENGHISMSTISVSEKFAAAHAEKGFLLAFLVLRPDAAKAAALRIITAGPAEKQA 120
CC
CC QY 413 CSSCFQAMGKTSFFLGE-----VGNAAKMLIVNVQGSFMAIAGLTLAHTVTCQSQOQL 469
CC : : : : : : : : : : : : : : : : : : : : : : : : : :

```



```

Db 121 AKPLDLSLQIIFDVCEESKTANA--ISINFLVSMLEALSESFLMKMKYGLEQKQFL 178
QY 470 DIINQOLASIFIDQKQCN-----ILQNFKP-DFYLYKYQKDLRLAIALGDVANHPTMA 524
Db 179 EI-----AKALFGSPYVGTITMAEQKPEPAGFKMSLGLKQNTNLALAAKRVSANPLA 233
QY 525 AANEVYKRAKALDQSDNDMSAVR 549
Db 234 ELAKSHFESGIEKGFGDLDAALIK 258

RESULT 17
Y770 MYCTU
ID Y770 MYCTU STANDARD; PRT; 295 AA.
AC P71825;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical oxidoreductase Rv0770/MT0794 (EC 1.1.1.-.-).
GN Rv0770 OR MT0794 OR MIC3369.15.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.F., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
RA "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490(2002).
CC CC
CC -!- SIMILARITY: Belongs to the 3-hydroxyisobutyrate dehydrogenase
family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; Z80226; CAB02382.1; -.
DR EMBL; AE006970; AAK45036.1; -.
DR PIR; E70707; E70707.
DR TIGR; MT0794; -.
DR TubercuList; Rv0770; -.
DR InterPro; IPR002204; 3hydroxisobut_dh.
DR PROSITE; PS00895; 3 HYDROXYISOBUT DH; 1.
DR Hypothetical protein; Oxidoreductase; NAD; Complete proteome.
FT ACT SITE 176 BY SIMILARITY.
SQ SEQUENCE 295 AA; 30445 MW; 7A9E7FBFAD0E95B8 CRC64;

```

```

Query Match 7.3%; Score 209.5; DB 1; Length 295;
Best Local Similarity 24.7%; Pred. No. 2.8e-06;
Matches 69; Conservative 48; Mismatches 137; Indels 25; Gaps 5;

QY 265 PDKKIGFLGLGMSGIVSNLLKMGHTVTVNRTAEKCDLFTQEGARLGRTPAEVVSTC 324
Db 5 PETPRLYIGLGNQAGAPMAKRLLDWPGGLTVFDRVVEAMAPFVEGGATAASVSD--VAEA 63
QY 325 DITFACVSDPKAAKDLVLGSPGVLOQIRPGKCVDMSTVDADTVELAQVIVSRGGRFLE 384
Db 64 DIISITVFDDAQVSSVITADNGLATHAKPGTIVATHSTIADTTAVDLAEKLPQGIHVD 123
QY 385 APVSGNQQLSNDGMLVILAAGDGLYEDCSCFCQAMGKTSFFLGEVGNAAKMLIVNVQ 444
Db 124 APVSGAAAKAGELAVVMVGADDEAFQRIKEPFSRWASLLIHAGBPGAGTRMKLARNMLT 183
QY 445 GSFMTIAEGLTIAHVTGSGQOQLDILNOGQL-----ASIFLDKQCNILQG 492
Db 184 FVSYAAMAEAQRLAEACG-----LDLVALGKVVRHSDSFTGGAGAIMFRNTTAPMEPAD 237
QY 493 NFKDPFYKYI-----QKDLRLAIALGDVANHPTMAAAA 527
Db 238 PLRP--LLEHTRGLGKDLALALGCVVVDLPLAQLA 274

RESULT 18
YGBJ HAEIN
ID YGBJ HAEIN STANDARD; PRT; 301 AA.
AC P44979;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical oxidoreductase Hli010 (EC 1.1.1.-.-).
GN Hli010.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RA "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RL Science 269:496-512(1995).
CC CC
CC -!- SIMILARITY: Belongs to the 3-hydroxyisobutyrate dehydrogenase
family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; U32782; AAC22671.1; -.
DR PIR; B64164; B64164.
DR TIGR; Hli010; -.
DR InterPro; IPR002204; 3hydroxisobut_dh.
DR InterPro; IPR006115; 6PGD NAD.
DR Pfam; PF03446; NAD binding 2; 1.
DR PROSITE; PS00895; 3 HYDROXYISOBUT DH; 1.
DR Hypothetical protein; Oxidoreductase; NAD; Complete proteome.
KW

```

Izumoto Y., Kuroda T., Harada H., Kishimoto T., Nakamura H.;
 RA "Hepato-ma-derived growth factor belongs to a gene family in mice
 RT showing significant homology in the amino terminus."; [PMID:10511000](#)
 RL Biochem. Biophys. Res. Commun. 238:26-32(1997).
 RN
 (2)
 RN SEQUENCE FROM N.A.
 RP STRAIN=BALE/c; TISSUE=Thymus;
 RC
 RA Zhao Y., Chen W., Wang Y.;
 RA "Cloning of novel gene related to thymus development."; [PMID:10511000](#)
 RT Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RL
 RN
 (3)
 RN SEQUENCE FROM N.A.
 RP STRAIN=C57BL/6J; TISSUE=Embryo;
 RC MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischnmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA

```

Qy 67 KEEMIKKNGKRRFOQAVDAVEFLRAKQKQDTSHNSDDKRRNSSEERSRPNSGDK 126
Db 70 KEKFGKPNKRGFSEGLWEIEN-----NPTVKAGYQSSQKSCAAAEPEVEPEAEGDGD 124
Qy 127 RKLISLSEGVKNWGEKGRVSSSSSSESGSKPLKR-----AQOSPRKGRPPK-----D 177
Db 125 KKGS-AEGSSDE--EGKLVIDEPAKEKNEKGTLLKRRAGDVLDESPPK---PRSEGDHE 176
Qy 178 EKDLTIPSSVTGMMAGPMAAFKQWQPTASBPVKDADP 215
Db 177 EED---KEIAALEGERPLPVEVEK-NSTPSPDGGQP 210

RESULT 20
YCY3 SHEFR
ID YCY3 SHEFR STANDARD; PRT; 194 AA.
AC O33730;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Hypothetical oxidoreductase in cytochrome c3 5' region (Fragment).
OS Shewanella frigidimarina.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
OX NCBI_TaxID=56812;
RN SEQUENCE FROM N.A.
RC STRAIN=NCIMB 400;
RX PubMed=10861223;
RA Gordon E.H.J., Pike A.D., Hill A.E., Cuthbertson P.M., Chapman S.K.,
RA Reid G.A.;
RT "Identification and characterization of a novel cytochrome c3 from
RT Shewanella frigidimarina that is involved in Fe(III) respiration.";
RL Biochem. J. 349:153-158(2000).
CC -i SIMILARITY: Belongs to the 3-hydroxyisobutyrate dehydrogenase
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AJ000006; CA03849.1; -
KW Hypothetical protein; Oxidoreductase; NAD.
FT NON_TER
SQ SEQUENCE 194 AA; 20865 MW; 4B4236749F1E8701 CRC64;

Query Match 6.5%; Score 185.5; DB 1; Length 194;
Best Local Similarity 26.3%; Pred. No. 5.2e-05;
Matches 51; Conservative 45; Mismatches 81; Indels 17; Gaps 5;

Qy 362 TVDADTVTELAQVIVSRGGRFLEAPVSGNQSLNDGMLVLAAGDRGLYE-----DCSS 415
Db 1 TAGADVAREIAAYEPLNIAFLDAPVSGQAGANGALTWMMGSDQAHFDTVKPVISAYS 60
Qy 416 CFQAMGKTSFELGVGNAAKMLVNMVQGSFMATIAEGLTIAHVTQSQOQLDILNQG 475
Db 61 CABELLPVG--AGQLTKRVNVQICITAGVVGQ-----LAEGHFAKSGLDGLKVEIVISKG 113
Qy 476 QLASIFLDQKONITLQCNFKPDEFVLKYIOKDLRLAIALGDAVNHPT--PMAAANAEVYKR 533
Db 114 RAQSWQENRYKTMQCCQYDFGFALDWRKD--LGIALDEARRNGSHLPVAALVDQFYSE 171
Qy 534 AKALDQSDNDMSAV 547
Db 172 VQAMKGNRWDTSLS 185

RESULT 21
HDGF_HUMAN

```

```

ID HDGF_HUMAN STANDARD; PRT; 240 AA.
AC P51858;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hepatoma-derived growth factor (HDGF) (High-mobility group protein 1-
DE like 2) (HMG-1L2).
GN HDGF OR HMG1L2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN SEQUENCE FROM N.A., AND SEQUENCE OF 4-24.
RC TISSUE=Hepatoma;
RX MEDLINE=95014294; PubMed=7929202;
RA Nakamura H., Izumoto Y., Kambe H., Kuroda T., Mori T., Kawamura K.,
RA Yamamoto H., Kishimoto T.;
RT "Molecular cloning of complementary DNA for a novel human hepatoma-
RT derived growth factor. Its homology with high mobility group-1
RT protein.";
RL J. Biol. Chem. 269:25143-25149(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.A., McEwan P.J., McKernan K.J., Abramson R.D., Mullahy S.J.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting R., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -i FUNCTION: Heparin-binding protein, with mitogenic activity for
CC fibroblasts.
CC -i SUBCELLULAR LOCATION: Cytoplasmic.
CC -i TISSUE SPECIFICITY: Ubiquitous.
CC -i SIMILARITY: Contains 1 PWMP domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: D16431; BAA03903.1; -
DR EMBL; BC018991; AH18991.1; -
DR PIR; A55055; A55055.
DR Genew; HGNC:4856; HDGF.
DR MIM; 300043; -
DR GO; GO:0005737; C:cytoplasm; TAS.
DR GO; GO:0005615; C:extracellular space; TAS.
DR GO; GO:0008201; F:heparin binding; TAS.
DR GO; GO:0008283; P:cell proliferation; TAS.
DR InterPro; IPR000313; PWMP_domain.
DR Pfam; PF00855; PWMP; 1.
DR SMART; SM00293; PWMP; 1.
DR PROSITE; PS50812; PWMP; 1.

```

KW	Growth factor; Heparin-binding.	
FT	DOMAIN 12 69	PWWP.
FT	DOMAIN 213 223	POLY-GLU.
SQ	SEQUENCE 240 AA; 26788 MW; DD60D9203BDD4B34 CRC64;	
	Query Match 6.5%; Score 185.5; DB 1; Length 240;	
	Best local Similarity 29.8%; Pred. No. 6.8e-05;	
	Matches 64; Conservative 30; Mismatches 100; Indels 21; Gaps 8	
QY	7 RLGLVWVKLGKRYPPWPKIVNPPKDLKKPRGKCKFFVKVFGFTGTHAWIKVQLKPYHAH 66	
DB	11 KCGDLVFAKMGYGFHWPARIDEMPEAAVKSTANK-YQVFFGTGTHAFLGPKDLFFYES 69	
QY	67 KEEMIKINKGRFQOAVDAVEEFLRAKQKQTSNHSDDKNRNSSEERSRPNNGDEK 126	
DB	70 KEKFGKPNKRKGFSEGLWEIEN-----NPTVKASGYQSSKQKSCVEEPEPEAAEGDGD 124	
QY	127 RKLSELSEGVKKNMGEGKKRVSSGSSERGSKPLKRAQ-----EQSPR--KGRPPPKDEKD 180	
DB	125 KK-GNAEGSSDE--EGKLAVIDEPAEKNEKRGALKRRAGDLDLDSFKRPKEAENPEGEE- 179	
QY	181 LTIPESSTVKMGMAAGPMAAFKQWPTASBPVKDADP 215	
DB	180 ---KEAATLEVERPLPMEVEK-NSTPSPGSGRGP 210	
RESULT 22		
MSH6 HUMAN		
ID	MSH6_HUMAN STANDARD; PRT; 1360 AA.	
AC	P52701; O43976; O43917; Q8TCX4;	
DT	01-OCT-1996 (Rel. 34, Created)	
DT	01-OCT-1996 (Rel. 34, Last sequence update)	
DT	15-MAR-2004 (Rel. 43, Last annotation update)	
DE	DNA mismatch repair protein MSH6 [MtrS-alpha 160 kDa subunit] (G/T	
DE	mismatch binding protein) (GTBP) (GTMBP) (P160).	
GN	MSH6 OR GTBP.	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OX	NCBI_TaxID=9606;	
[1]		
RP	SEQUENCE FROM N.A.	
RP	MEDLINE=97098445; PubMed=8942385;	
RA	Acharya S., Wilson T., Gradia S., Kane M.F., Guerrette S.,	
RA	Marsischky G.T., Kolodner R.D., Fishel R.;	
RT	"hMSH2 forms specific mismatch binding complexes with hMSH3 and	
RT	hMSH6.";	
RL	Proc. Natl. Acad. Sci. U.S.A. 93:13629-13634 (1996).	
[2]		
RP	SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.	
RP	MEDLINE=98116665; PubMed=9455487;	
RA	Shiwaku H.O., Wakatsuki S., Mori Y., Fukushige S., Horii A.;	
RT	"Alternative splicing of GTBP in normal human tissues.";	
RL	DNA Res. 4:359-362 (1997).	
[3]		
RP	SEQUENCE FROM N.A., AND VARIANT GLY-39.	
RP	Rieder M.J., Braun A.C., Montoya M.A., Chung M.-W., Nguyen C.P.,	
RA	Nguyen D.A., Livingston R.J., Poel C.L., Robertson P.D.,	
RA	Schnackwitz W.S., Sherwood J.K., Witrak I.A., Nickerson D.A.;	
RL	Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.	
[4]		
RP	SEQUENCE OF 69-1360 FROM N.A., AND PARTIAL SEQUENCE.	
RP	MEDLINE=95327934; PubMed=7604265;	
RA	Palombo F., Gallinari P., Iaccarino I., Lettieri T., Hughes M.,	
RA	D'Arrigo A., Truong O., Hsuan J.J., Jiricny J.;	
RT	"GTBP, a 160-kilodalton protein essential for mismatch-binding	
RT	activity in human cells.";	
RL	Science 268:1912-1914 (1995).	
[5]		
RP	SEQUENCE OF 1-116 FROM N.A., AND VARIANT GLY-39.	
RP	MEDLINE=96435440; PubMed=8838326;	
RA	Nicolaidis N.C., Palombo F., Kinzler K.W., Vogelstein B.,	
RA	Jiricny J.;	

RA Schackert H. K. ;
 RT "Involvement of hMSH6 in the development of hereditary and sporadic
 RT colorectal cancer revealed by immunostaining is based on germline
 RT mutations, but rarely on somatic inactivation." ;
 RL Int. J. Cancer 97:643-648(2002).
 CC -!- FUNCTION: Restores repair of base-base and single- nucleotide
 CC insertion-deletion mismatches, and increases the proficiency to
 CC process heteroduplexes with two-, three-, or four- nucleotide
 CC insertion-deletion mismatches. GTBP binds covalently to G/T
 CC mismatches.
 CC -!- SUBUNIT: Heterodimer of MSH2 and MSH6 (GTBP). Part of the BRCA1-
 CC associated genome surveillance complex (BASC), which contains
 CC BRCA1, MSH2, MSH6, MLH1, ATM, BLM, PMS2 and the RAD50-MRE11-NBS1
 CC protein complex. This association could be a dynamic process
 CC changing throughout the cell cycle and within subnuclear domains
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=alternative splicing; Named isoforms=2;
 CC Name=GTBP-N;
 CC IsoId=P52701-1; Sequence=Displayed;
 CC Name=GTBP-alt;
 CC IsoId=P52701-2; Sequence=VSP 003291, VSP_003292;
 CC -!- PM: The N-terminus is blocked"
 CC -!- DISEASE: Defects in MSH6 are the cause of hereditary non-polyposis
 CC colorectal cancer type 5 (HNPPC5) [MIM:600678]. Mutations in more
 CC than one gene locus can be involved alone or in combination in the
 CC production of the HNPPC phenotype (also called Lynch syndrome).
 CC Most families with clinically recognized HNPPC have mutations in
 CC either MLH1 or MSH2 genes. HNPPC is an autosomal, dominantly
 CC inherited disease associated with marked increase in cancer
 CC susceptibility. It is characterized by a familial predisposition
 CC to early onset colorectal carcinoma (CRC) and extra-colonic
 CC cancers of the gastrointestinal, urological and female
 CC reproductive tracts. HNPPC is reported to be the most common form
 CC of inherited colorectal cancer in the Western world. Cancers in
 CC HNPPC originate within benign neoplastic polyps termed adenomas.
 CC Clinically, HNPPC is often divided into two subgroups. Type I:
 CC hereditary predisposition to colorectal cancer, a young age of
 CC onset, and carcinoma observed in the proximal colon. Type II:
 CC patients have an increased risk for cancers in certain tissues
 CC such as the uterus, ovary, breast, stomach, small intestine, skin,
 CC and larynx in addition to the colon. Diagnosis of classical HNPPC
 CC is based on the Amsterdam criteria: 3 or more relatives affected
 CC by colorectal cancer, one a first degree relative of the other
 CC two; 2 or more generation affected; 1 or more colorectal cancers
 CC presenting before 50 years of age; exclusion of hereditary
 CC polyposis syndromes. MSH6 mutations appear to be associated with
 CC atypical HNPPC and in particular with development of endometrial
 CC carcinoma or atypical endometrial hyperplasia, the presumed
 CC precursor of endometrial cancer. Defects in MSH6 are also found in
 CC familial colorectal cancers (suspected or incomplete HNPPC) that
 CC do not fulfill the Amsterdam criteria for HNPPC.
 CC -!- DISBAS: Defects in MSH6 are a cause of susceptibility to
 CC endometrial cancer [MIM:608089].
 CC -!- SIMILARITY: Belongs to the DNA mismatch repair muts family.
 CC -!- SIMILARITY: Contains 1 PWWP domain.
 CC -!- DATABASE: NAME=Hereditary non-polyposis colorectal cancer db;
 CC WWW="http://www.nfdrift.nl/".

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR	EMBL; U73737; AAB47425.1; -
DR	EMBL; U73732; AAB47425.1; JOINED.
DR	EMBL; U73733; AAB47425.1; JOINED.
DR	EMBL; U73734; AAB47425.1; JOINED.
DR	EMBL; U73736; AAB47425.1; JOINED.
DR	EMBL; D89645; BAA23674.1; JOINED.

DR	EMBL	D89637	BAA23674.1	JOINED.
DR	EMBL	D89638	BAA23674.1	JOINED.
DR	EMBL	D89639	BAA23674.1	JOINED.
DR	EMBL	D89640	BAA23674.1	JOINED.
DR	EMBL	D89641	BAA23674.1	JOINED.
DR	EMBL	D89642	BAA23674.1	JOINED.
DR	EMBL	D89643	BAA23674.1	JOINED.
DR	EMBL	D89644	BAA23674.1	JOINED.
DR	EMBL	D89645	BAA23675.1	-.
DR	EMBL	D89646	BAA23675.1	-.
DR	EMBL	D89647	BAA23676.1	-.
DR	EMBL	D89648	BAA23677.1	-.
DR	EMBL	D89649	BAA23678.1	-.
DR	EMBL	D89650	BAA23679.1	-.
DR	EMBL	D89651	BAA23680.1	-.
DR	EMBL	D89652	BAA23681.1	-.
DR	EMBL	D89653	BAA23682.1	-.
DR	EMBL	D89654	BAA23683.1	-.
DR	EMBL	D89655	BAA23684.1	-.
DR	EMBL	D89656	BAA23685.1	-.
DR	EMBL	D89657	BAA23686.1	-.
DR	EMBL	D89658	BAA23687.1	-.
DR	EMBL	D89659	BAA23688.1	-.
DR	EMBL	D89660	BAA23689.1	-.
DR	EMBL	D89661	BAA23690.1	-.
DR	EMBL	D89662	BAA23691.1	-.
DR	EMBL	D89663	BAA23692.1	-.
DR	EMBL	D89664	BAA23693.1	-.
DR	EMBL	D89665	BAA23694.1	-.
DR	EMBL	D89666	BAA23695.1	-.
DR	EMBL	D89667	BAA23696.1	-.
DR	EMBL	D89668	BAA23697.1	-.
DR	EMBL	D89669	BAA23698.1	-.
DR	EMBL	D89670	BAA23699.1	-.
DR	EMBL	D89671	BAA23700.1	-.
DR	EMBL	D89672	BAA23701.1	-.
DR	EMBL	D89673	BAA23702.1	-.
DR	EMBL	D89674	BAA23703.1	-.
DR	EMBL	D89675	BAA23704.1	-.
DR	EMBL	D89676	BAA23705.1	-.
DR	EMBL	D89677	BAA23706.1	-.
DR	EMBL	D89678	BAA23707.1	-.
DR	EMBL	D89679	BAA23708.1	-.
DR	EMBL	D89680	BAA23709.1	-.
DR	EMBL	D89681	BAA23710.1	-.
DR	EMBL	D89682	BAA23711.1	-.
DR	EMBL	D89683	BAA23712.1	-.
DR	EMBL	D89684	BAA23713.1	-.
DR	EMBL	D89685	BAA23714.1	-.
DR	EMBL	D89686	BAA23715.1	-.
DR	EMBL	D89687	BAA23716.1	-.
DR	EMBL	D89688	BAA23717.1	-.
DR	EMBL	D89689	BAA23718.1	-.
DR	EMBL	D89690	BAA23719.1	-.
DR	EMBL	D89691	BAA23720.1	-.
DR	EMBL	D89692	BAA23721.1	-.
DR	EMBL	D89693	BAA23722.1	-.
DR	EMBL	D89694	BAA23723.1	-.
DR	EMBL	D89695	BAA23724.1	-.
DR	EMBL	D89696	BAA23725.1	-.
DR	EMBL	D89697	BAA23726.1	-.
DR	EMBL	D89698	BAA23727.1	-.
DR	EMBL	D89699	BAA23728.1	-.
DR	EMBL	D89700	BAA23729.1	-.
DR	EMBL	D89701	BAA23730.1	-.
DR	EMBL	D89702	BAA23731.1	-.
DR	EMBL	D89703	BAA23732.1	-.
DR	EMBL	D89704	BAA23733.1	-.
DR	EMBL	D89705	BAA23734.1	-.
DR	EMBL	D89706	BAA23735.1	-.
DR	EMBL	D89707	BAA23736.1	-.
DR	EMBL	D89708	BAA23737.1	-.
DR	EMBL	D89709	BAA23738.1	-.
DR	EMBL	D89710	BAA23739.1	-.
DR	EMBL	D89711	BAA23740.1	-.
DR	EMBL	D89712	BAA23741.1	-.
DR	EMBL	D89713	BAA23742.1	-.
DR	EMBL	D89714	BAA23743.1	-.
DR	EMBL	D89715	BAA23744.1	-.
DR	EMBL	D89716	BAA23745.1	-.
DR	EMBL	D89717	BAA23746.1	-.
DR	EMBL	D89718	BAA23747.1	-.
DR	EMBL	D89719	BAA23748.1	-.
DR	EMBL	D89720	BAA23749.1	-.
DR	EMBL	D89721	BAA23750.1	-.
DR	EMBL	D89722	BAA23751.1	-.
DR	EMBL	D89723	BAA23752.1	-.
DR	EMBL	D89724	BAA23753.1	-.
DR	EMBL			

RESULT 23
6PG9 ECOI.T

ID	PG9_ECOLI	STANDARD;	PRT;	468 AA.
AC	P37754;			
DT	01-OCT-1994	(Rel. 30, Created)		
DT	01-OCT-1994	(Rel. 30, Last sequence update)		
DT	28-FEB-2003	(Rel. 41, Last annotation update)		
DE	6-phosphogluconate dehydrogenase, decarboxyl			
GN	GND.			
OS	Escherichia coli.			

DR pfam; PF00855; PWWP; 1.
 DR ProDom; PD001263; Muts_C; 1.
 DR SMART; SM00534; MUTSac; 1.
 DR SMART; SM00533; MUTSG; 1.
 DR SMART; SM00293; PWWP; 1.
 DR PROSITE; PS00486; DNA_MISMATCH_REPAIR_2; 1.
 DR PROSITE; PS00812; PWWP; 1.
 KW DNA repair; ATP-binding; DNA-binding.
 FT DOXAIN 92 154 PWWP.
 FT DOMAIN 201 209 POLY-GLU.
 FT BIND 1132 1139 ATP (POTENTIAL).
 FT CONFLICT 65 66 DG -> EA (IN REF. 2).
 FT CONFLICT 374 375 OK -> PE (IN REF. 2).
 FT CONFLICT 754 754 N -> T (IN REF. 2).
 FT CONFLICT 800 800 Y -> D (IN REF. 2).
 FT CONFLICT 1227 1227 S -> N (IN REF. 2).
 FT CONFLICT 1329 1329 R -> Q (IN REF. 2).
 FT CONFLICT 1333 1333 G -> E (IN REF. 2).
 SQ SEQUENCE 1358 AA; 151075 MW; 2031F78D477804AD CRC64;

Query Match 5.1%; Score 145.5; DB 1; Length 1358;
 Best Local Similarity 24.0%; Pred. No. 0.22; Mismatches 29; Conservative 49; Indels 89; Gaps 7;
 Matches 49; Conservative 29; Mismatches 29; Indels 89; Gaps 7;

QY 9 GDLVWGKLGKRYPPWPKIVNPPKD--LKKPRGKCC-FVVKPF-GTETHAWIKVEQLKPYH 64
 DB 93 GDLVWAKMEGYPPWPCLYVNHPPFDGTFIRKKGKSVRVHVQFFDDSPTRGWVSKRMKLPYT 152
 QY 65 AHKEEMIKINKGRFQQAQVDAVERFLRA-----KQKDTSS 101
 DB 153 GSKSK--EAQKGHFYSKSEILRAMORADALSKDTAERLQLAVCDPEPSEEEETE 210
 QY 102 H----NSSDDKNRNSSEERSRPNSSGDEKRLKLSSEGKVKKNMGEGKKRVSSGSSRGSK 157
 DB 211 HEAVLSKSEEDNYSSEEAQPSVQGR-----SSRQVKKRVISDESIDIGSSDVEFK 266
 QY 158 SPLKRAQEQSKRKRPPKDEKDL 181
 DB 267 PDTKQEGSDDDASSGVGDSSEDL 290

RESULT 25
 6PGD CUNEL
 ID 6PGD CUNEL STANDARD; PRT; 485 AA.
 AC O60037;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44).
 GN 6-PGD.
 OS Cunninghamella elegans.
 OC Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales;
 OC Cunninghamellaceae; Cunninghamella.
 OX NCBI_TaxID=4853;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 36112;
 RX MEDLINE=99085678; PubMed=9868787;
 RA Wang R.F., Khan A.A., Cao W.W., Cerniglia C.E.;
 RT "Identification and sequencing of a cDNA encoding 6-phosphogluconate
 RT dehydrogenase from a fungus, Cunninghamella elegans and expression of
 RT the gene in Escherichia coli.";
 RL FEMS Microbiol. Lett. 169:397-402(1998).
 CC -|- CATALYTIC ACTIVITY: 6-phospho-D-gluconate + NADP(+) = D-ribulose
 CC 5-phosphate + CO(2) + NADPH.
 CC -|- PATHWAY: Hexose monophosphate shunt.
 CC -|- SIMILARITY: Belongs to the 6-phosphogluconate dehydrogenase
 CC family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Y17297; CAA76734.1; .
 DR HSSP; P00349; 2PGD.
 DR InterPro; IPR008927; 6DGDH_C-like.
 DR InterPro; IPR006183; 6PGD_C.
 DR InterPro; IPR006114; 6PGD_C.
 DR InterPro; IPR006113; 6PGD decarbox.
 DR InterPro; IPR006115; 6PGD NAD.
 DR InterPro; IPR006184; 6PGD_BS.
 DR Pfam; PF00393; 6PGD; 1.
 DR Pfam; PF03446; NAD binding 2; 1.
 DR PRINTS; PR00076; 6PGDHDRGNASE.
 DR TIGRFAMs; TIGR00873; gnd; 1.
 DR PROSITE; PS00461; 6PGD; 1.
 KW Oxidoreductase; Pentose shunt; NADP.
 SQ SEQUENCE 485 AA; 53102 MW; F7E6DDFF3D21EFB2 CRC64;

Query Match 5.0%; Score 143.5; DB 1; Length 485;
 Best Local Similarity 22.7%; Pred. No. 0.074;
 Matches 66; Conservative 50; Mismatches 140; Indels 35; Gaps 9;

QY 270 IGFGLGLMGSGIYSNLLKMGHTVTVNRTAEKCDLFIQEGARLCRTPAEVVSTCDITFA 329
 DB 8 IGLGLAVGQQLINNDHGVVCAYNRITSKVDDFLANEAK---GTNVVGAHSVEEL 63
 QY 330 CVSDPKAAKDILVLPSPG-----VLQIRPGKCVDMSTVDADTVELAQVIVSRGG 380
 DB 64 CAKLKPRKVMLLVKAGSAVDADFIDQLLPLEEGDIIDGNSHFPDSIRTKLEAKGI 123
 QY 381 RFLAPVSGNQQLSNDGMVLILAGDRLGVDCSCFQAMGKTS-----FFLGEVGN 433
 DB 124 LFGVSGVSGGEGARYGP-SLMPGSGSKAWEHIQPIQATAAKADGASCCEWVGETGAG 182
 QY 434 AKOMLIVNVQSGSPWATIAEGLTLAHV-TGQSQTLDLIL---NQQLASIFLQKQNI 489
 DB 183 HYVKNVHNGIEYGDMLITEVYQILHEGLGSHDEMADIFEWNKGDLDS-FLIEITRDI 241
 QY 490 LQGNFKPDPYLVKIQKDLRLAIALG-----DAVNHTPMAAAANAEVYKR 533
 DB 242 LR--FKTDGQPLVTKIRDTAGQKGTGKTALDLSLDRGIPVTLIGEAVYSR 290

RESULT 26
 6PGD KLEPN
 ID 6PGD KLEPN STANDARD; PRT; 468 AA.
 AC P41576; Q48461;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44).
 GN GND.
 OS Klebsiella pneumoniae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Klebsiella.
 OX NCBI_TaxID=573;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Chedid;
 RX MEDLINE=95204345; PubMed=7896702;
 RA Arakawa Y., Wacharotayankun R., Nagatsuka T., Ito H., Kato N.,
 RA Ohta M.;
 RT "Genomic organization of the Klebsiella pneumoniae cps region
 RT responsible for serotype K2 capsular polysaccharide synthesis in the
 RT virulent strain Chedid".
 RL J. Bacteriol. 177:1788-1796(1995).
 RN [2]
 RP SEQUENCE OF 12-456 FROM N.A.
 RC STRAIN=CW 7380;
 RX MEDLINE=95024018; PubMed=7937867;
 RA Nelson K., Selander R.K.,


```

RT "Intergenic transfer and recombination of the 6-phosphogluconate
RT dehydrogenase gene (gnd) in enteric bacteria.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:10227-10231(1994).
CC -!- CATALYTIC ACTIVITY: 6-phospho-D-gluconate + NADP(+) = D-ribulose
CC 5-phosphate + CO(2) + NADPH.
CC -!- PATHWAY: Hexose monophosphate shunt.
CC -!- SIMILARITY: Belongs to the 6-phosphogluconate dehydrogenase
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DE EMBL; D21242; BRA04786.1; -
DR EMBL; U14471; AAC43817.1; -
DR PIR; D56146; D56146.
DR HSSP; P00349; 2PGD.
DR InterPro; IPR008927; 6GDGH_C_like.
DR InterPro; IPR006183; 6PGD.
DR InterPro; IPR006114; 6PGD_C.
DR InterPro; IPR006113; 6PGD-decarbox.
DR InterPro; IPR006115; 6PGD_NAD.
DR Pfam; PF00393; 6PGD; 1.
DR Pfam; PF03446; NAD binding 2; 1.
DR PRINTS; PR00076; 6EGDHDGRNASE.
DR TIGRfams; TIGR00873; gnd; 1.
DR PROSITE; PS00461; 6PGD; 1.
KW Gluconate utilization; Oxidoreductase; Pentose shunt; NADP.
FT CONFLICT 316 316 G -> E (IN REF. 2).
FT CONFLICT 421 421 V -> F (IN REF. 2).
SQ SEQUENCE 468 AA; 51328 MW; FF1EB5E765FDC90 CRC64;

Query Match 4.9%; Score 141; DB 1; Length 468;
Best Local Similarity 21.4%; Pred. No. 0.1;
Matches 68; Conservative 51; Mismatches 113; Indels 86; Gaps 12;

QY 268 KTGFLGLGMSGIVSNLLKMGHTVTVMNRATKCDLFTQCARLGRTPAEVSFCDIT 327
DB 4 QQLGVGVMAVMGNLALNTESRGTYTVSVNRSREKTEEVIAENTGKLVYTVQV---F 60
QY 328 FACVSDPKAAKDLVGLPGSGVLOGIRPKCYVMSTVDVDTVTELAQVIVSRGRFLEAPV 387
DB 61 VESLETPRILLVMKAGAGTSDAISLKPYLDK-----GDIIIDGGNTFFQDTI 109
QY 388 SGNQQLSNDGMLVI---LAAGDRG-----LYEDCSSC 416
DB 110 RRRNELSAEFGNFIGTVSGSGEGALKGPSIMPGQKEAYELVAPILKQIAVAEAGEPC 169
QY 417 FQAMGKTSFPLGEVGNAAKMLLVMVQGSFMATIAE-----GLTLAHTVQSQQTLL 469
DB 170 VT-----YIGADGAGHYVMVHNGIEYGDMLIAEYALLKGLALS- -BELAQTF 220
QY 470 DILNQGLASIFLDKQNTLOQNF-XPDPYLKVIQKDLRLAIALGDAVNHPT----- 521
DB 221 E-WNEGELSYLID-----ITKDIFTKKDEGKYL-----VDVILDEAANKGTRKTSQS 269
QY 522 -----PMAAAANEVYKR 533
DB 270 SLDLGEPLSLITESVFAR 287

RESULT 27
ID 6PGD_SCHPO STANDARD; PRT; 492 AA.
AC P78812; Q9UQW5;
DT 15-JUL-1998 (Rel. 36, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

```

```

DE GN 6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44).
OS SPBC660.16.
OS Schizosaccharomyces pombe (Fission Yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PR745;
RX MEDLINE=98162722; PubMed=9501991;
RA Yoshiooka S., Kato K., Nakai K., Okayama H., Nojima H.;
RT "Identification of open reading frames in Schizosaccharomyces pombe
RT CDNAS.";
RL DNA Res. 4:363-369(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser G.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson A.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Stevens K.,
RA Skelton J., Simmonds M., Squares R., Squares S., Warren T., Whitehead S.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Cabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B., Mottier S.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Lelaure V.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armetrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -!- CATALYTIC ACTIVITY: 6-phospho-D-gluconate + NADP(+) = D-ribulose
CC 5-phosphate + CO(2) + NADPH.
CC -!- PATHWAY: Hexose monophosphate shunt.
CC -!- SIMILARITY: Belongs to the 6-phosphogluconate dehydrogenase
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DE EMBL; D89161; BAA13823.1; AUT_INIT.
DR EMBL; AL034563; CAA22536.1; -.
DR PIR; T40628; T40628.
DR HSSP; P00349; 2PGD.
DR GenedB SPombe; SPBC660.16; -.
DR InterPro; IPR008927; 6GDGH_C_like.
DR InterPro; IPR006183; 6PGD.
DR InterPro; IPR006114; 6PGD_C.
DR InterPro; IPR006113; 6PGD-decarbox.
DR InterPro; IPR006115; 6PGD_NAD.
DR InterPro; IPR006184; 6PGDm_BS.
DR Pfam; PF00393; 6PGD; 1.
DR Pfam; PF03446; NAD binding 2; 1.
DR PRINTS; PR00076; 6EGDHDGRNASE.

```

```

DR TIGRFAMS; TIGR00873; gnd; 1.
DR PROSITE; PS00461; 6PGD; 1.
KW Oxidoreductase; Pentose shunt; NADP.
FT CONFLICT 219 220 IA -> ST (IN REF. 1).
SQ SEQUENCE 492 AA; 53679 MW; F55F342957A9D3E1 CRC64;

Query Match
Best Local Similarity 4.8%; Score 138.5; DB 1; Length 492;
Matches 68; Conservative 53; Mismatches 127; Indels 49; Gaps 12;

QY 271 GFLGLGMLSGIVSNLLKMGHTVTVNRTAEKCDLFIQEGARLGRTPA-----EVSVC 324
Db 10 GLIGLAVMGQNLINGADKGFTVCVYNTTSTVDFLANEAK-GRKSIIVGAHSLSEFVSKL 68

QY 325 DITFACVSDPKAAKDLVLGSPGVLOGIRP---GKCYVDMSTVDADTVELAQVIVSRGG 380
Db 69 KKPVCILLVRAGKEV---DVLIEGLAPLEKGIIVDGGNSHYPTDTRCEEELAKKGI 124

QY 381 RPLEAPVSGNQLSNDGMVLVLAAGDRGLYEDCSFCQAMGKTS-----FFLGEVGN 433
Db 125 LFGVSGVSGEGARYGP-SLMPGNGPAAPRIKPIFOTLAAKAGNNEPCCDWVGEQAG 183

QY 434 AKMMLIVNMQGSEFMATAE-----GLTLAHVTGQSQOITLIDL---NQQLASIFLD 483
Db 184 HYKMHVHNGIEYGDNLICETVDIMKRGD-----GMSCEIADVFKEKWTGKLDL-S 236

QY 484 QKQNILOGNFKPDPFLYKYIQKDLRLALALG-----DAVNHPTPMAAANEVYKR 533
Db 237 EITRDVLR--YKADGKPLVEKILDAAGQKGTGKTAQNALEMGTPVSLITEAVFAR 291

RESULT 28
ID 6PGD SHIFL STANDARD; PRT; 468 AA.
AC P37756;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE 6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44).
GN GND OR SF2091 OR S2212.
OS Shigella flexneri
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
[1]
SEQUENCE FROM N.A.
STRAIN=PE577 / Serotype 2a;
MEDLINE=22272406; PubMed=12384590;
RX Yang J., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
through comparison with genomes of Escherichia coli K12 and O157."
RL Nucleic Acids Res. 30:4432-4441(2002).
[3]
SEQUENCE FROM N.A.
STRAIN=2457T / Arcc 700930 / Serotype 2a;
MEDLINE=22590274; PubMed=12704152;
RX Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
flexneri serotype 2a strain 2457T."

```

RESULT 29

```

RL Infect. Immun. 71:2775-2786(2003).
RN [4]
RP SEQUENCE OF 12-456 FROM N.A.
RC STRAIN=ATCC 29903;
RX MEDLINE=95024018; PubMed=7937867;
RA Nelson K., Selander R.K.;
RT "Intergeneric transfer and recombination of the 6-phosphogluconate
dehydrogenase gene (gnd) in enteric bacteria."
RL Proc. Natl. Acad. Sci. U.S.A. 91:10227-10231(1994).
CC -!- CATALYTIC ACTIVITY: 6-phospho-D-gluconate + NADP(+) = D-ribulose
5-phosphate + CO(2) + NADPH.
CC -!- PATHWAY: Hexose monophosphate shunt.
CC -!- SIMILARITY: Belongs to the 6-phosphogluconate dehydrogenase
family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
DR EMBL; X71970; CAA50781.1; -.
DR EMBL; AE015225; AAN43630.1; -.
DR EMBL; AE016985; AAP17458.1; -.
DR EMBL; U14468; AAC43834.1; -.
DR HSP; P00349; 2PGD.
DR InterPro; IPR008927; 6PGDH_C like.
DR InterPro; IPR006183; 6PGD.
DR InterPro; IPR006114; 6PGD_C.
DR InterPro; IPR006113; 6PGD-decarbox.
DR InterPro; IPR006115; 6PGD_NAD.
DR InterPro; IPR006184; 6PGDm_BS.
DR Pfam; PF00393; 6PGD; 1.
DR Pfam; PF03446; NAD binding 2; 1.
DR PRINTS; PR00076; 6PGDHRGNASE.
DR TIGRFAMS; TIGR00873; gnd; 1.
DR PROSITE; PS00461; 6PGD; 1.
KW Gluconate utilization; Oxidoreductase; Pentose shunt; NADP;
Complete proteome.
SQ SEQUENCE 468 AA; 51344 MW; 2203A0D82120CD61 CRC64;

Query Match 4.8%; Score 137; DB 1; Length 468;
Best Local Similarity 20.7%; Pred. No. 0.18;
Matches 68; Conservative 54; Mismatches 100; Indels 106; Gaps 14;

QY 268 KKIGFLGLGMLSGIVSNLLKMGHTVTVNRTAEKCDLFIQEGARLGRTPA-----EVSVC 323
Db 4 QQIGVGVGMAVGNRLALNIESRGYTVSIFNRSREKTEVIAENPGKKLAPYTVKEFVES 63

QY 324 CDITFACVSDPKAAKDLVLGSPGVLOGIRPDKCYVDMSTVDADTVELAQVIVSRGRFL 383
Db 64 -----LETPRILLMVKAGAGTDAIDSLKPYLDK-----GDIIDGNTFF 105

QY 384 EAPVSGNQQLSNDGMVLVLAAGDRG-----LYED 412
Db 106 QDITRRNRRLSAGFNFIQTVSGGEGALKGPSIMFGQKEAYELVAPILTKIAVAED 165

QY 413 CSSCFQAMGKTSFFLGEVGNAAKMLIVNMQGSEFMATAE-----GLTLAHVTGQSQ 465
Db 166 GEPVTV-----YIGADGAGHYVKNVHNGIEYGDMLIAEAYSLKGGNLNSW--BELA 216

QY 466 QTLIDLINQQLASIFLD-----QKQNILOGNFKPDPFLYKYIQKDLRLALALGDAVN 518
Db 217 QTTFE-WNGELSSYLIDITKIDFTKDED---GNLVLD-----VILDEAN 259

QY 519 HPT-----PMAAANEVYKR 533
Db 260 KGTGKWTQSALDGLGEPLSLITESVFAR 287

```

[illegible]

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Balwani R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brattier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dou L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Folsler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiland T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K.A., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RA "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 [2]
 RP REVISIONS.
 RX MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Bettencourt B.R., Celisner S.E., de Grey A.D.N.J., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
 RT systematic review.";
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley; TISSUE=Embryo;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.W.,
 RA Champagne M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA George R.A., Gonzalez M., Guarin H., Krommiller B., Li P.W., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Pacle J.M., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K.H., Yu C., Lewis S.E., Rubin G.M.,
 RA Celisner S.E.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Potential histone methyltransferase. Histone methylation
 CC gives specific tags for epigenetic transcriptional activation or
 CC repression (by similarity).
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- SIMILARITY: Contains 3 PHD-type zinc fingers.
 CC -!- SIMILARITY: Contains 1 post-SET domain.
 CC -!- SIMILARITY: Contains 2 PWMP domains.
 CC -!- SIMILARITY: Contains 1 SET domain.
 CC -!- CAUTION: Ref.3 sequence differs from that shown due to a
 CC frameshift in position 348.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC -----
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE003763; AAF56762.2; -;
 DR EMBL; AY118404; AAM48433.1; ALT_FRAME.
 DR FlyBase; FBGN0039559; Mes-4.
 DR GO; GO:0000228; C:nuclear chromosome; ISS.
 DR GO; GO:0042054; F:histone methyltransferase activity; ISS.
 DR GO; GO:0016458; P:gene silencing; ISS.
 DR GO; GO:0018992; P:germ-line sex determination; ISS.
 DR InterPro; IPR006560; AWS.
 DR InterPro; IPR003616; PostSET.
 DR InterPro; IPR000313; PWMP_domain.
 DR InterPro; IPR001214; SET.
 DR InterPro; IPR001965; Znf_PHD.
 DR Pfam; PF00628; PHD; 1.
 DR Pfam; PF00855; PWMP; 1.
 DR Pfam; PF00856; SET; 1.
 DR SMART; SM00570; AWS; 1.
 DR SMART; SM00249; PHD; 3.
 DR SMART; SM00293; PWMP; 2.
 DR SMART; SM00317; SET; 1.
 DR PROSITE; PS00868; POST_SET; 1.
 DR PROSITE; PS00812; PWMP_2.
 DR PROSITE; PS0280; SET; 1.
 DR PROSITE; PS01359; ZF_PHD_1; 1.
 DR PROSITE; PS0016; ZF_PHD_2; 1.
 KW Nuclear protein; zinc-finger; Metal-binding; Repeat.
 FT DOMAIN 395 468
 FT ZN_FING 777 862
 FT ZN_FING 864 932
 FT ZN_FING 999 1044
 FT DOMAIN 1049 1111
 FT DOMAIN 1233 1355
 FT DOMAIN 1359 1375
 FT DOMAIN 95 100
 FT DOMAIN 174 181
 SQ SEQUENCE 1427 AA; 159027 MW; FB6EFD62E814BEF3 CRC64;
 Query Match 4.7%; Score 135.5; DB 1; Length 1427;
 Best Local Similarity 30.7%; Pred. No. 0.98;
 Matches 35; Conservative 16; Mismatches 40; Indels 23; Gaps 5;
 QY 9 GDLVWGKLGYPWPWGKIVNPKD-----LKKPRGKCKFFVKFFGTEDHAIKVEQLKPY 63
 Db 1050 GEIVWAKFNFRWPA-IILPPTVEVPSNLIKKAHGENDFVVRFFGTHDHWISRRVLY 1108
 QY 64 -----HAHKEEMIKINKKFKQAVDAVEFLR-----RAKGDQDTSSHNSD 106
 Db 1109 IEGTGTGCHK-----TKSQLFRNYYTTCVEASRFLPIIKARQEQDMEROSGN 1156
 Search completed: September 16, 2004, 07:30:47
 Job time : 17 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2004, 07:28:53 ; Search time 48 Seconds
(without alignments)
3635.033 Million cell updates/sec

Title: US-10-067-482-2
Perfect score: 2866
Sequence: 1 MAASVLRGLDVLWGKGRYP.....AKALDQSDNDSAVYRAYIH 553

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database :

- 1: SP TREMBL 25.*
- 2: sp archaea.*
- 3: sp bacteria.*
- 4: sp fungi.*
- 5: sp human.*
- 6: sp invertebrate.*
- 7: sp mammal.*
- 8: sp mic.*
- 9: sp organelle.*
- 10: sp phase.*
- 11: sp plant.*
- 12: sp rodent.*
- 13: sp virus.*
- 14: sp vertebrate.*
- 15: sp unclassified.*
- 16: sp rvirus.*
- 17: sp bacteriap.*
- 18: sp archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2811	98.1	547	4	Q9BXX2
2	2767.5	96.6	546	11	Q922P9
3	2745.5	95.8	545	11	Q9CYQ1
4	2683.5	93.6	523	4	Q9BT11
5	796.5	27.8	602	5	Q8T079
6	727	25.4	293	10	Q84VC8
7	697	24.3	289	10	Q94B07
8	695	24.2	289	10	Q9LSV0
9	667	23.3	343	10	Q8RWF1
10	644	22.5	387	10	Q9LNP0
11	562	19.6	248	10	Q8LQJ7
12	560	19.5	343	10	Q8LQJ7
13	397	13.9	300	17	Q8TT25
14	392.5	13.7	288	16	O66454
15	392	13.7	289	17	Q97XZ7
16	384	13.4	334	10	Q9SZE1

Q97DX8 clostridium
Q8ZXR3 pyrobaculum
Q89NA0 bradyrhizob
Q8U2W2 pyrococcus
Q8EDH8 shewanella
Q8XAE4 escherichia
Q8ZLV8 salmonella
Q819E3 bacillus ce
Q8ZK1 salmonella
Q83K1 streptomyce
Q83Q26 shigella fl
Q7WFP7 bordetella
Q8ESP3 oceanobacil
Q83A7 enterococcu
Q7W489 bordetella
Q82MW0 streptomyce
Q8BUE9 lactobacill
Q8D3Q1 vibrio vuln
Q8M7 rhizobium l
Q89RT2 bradyrhizob
Q81Q70 oryza sativ
Q7UTL3 rhodopirell
Q8I20 rhizobium l
Q87G98 vibrio para
Q8TX6 pseudomonas
Q89Y2 brucella su
Q8YGI1 brucella me
Q8VP31 escherichia
Q7WG50 bordetella
Q8Z597 streptomyce
Q82D17 listeria in
Q8XNE4 clostridium
Q81DR6 thizobium m
Q81DR6 bacillus ce
Q9K911 listeria ha
Q8Y8A5 listeria mo
Q82L15 streptomyce
Q8ZLJ4 helicobacte
Q8YJH6 neisseria m
Q8YTH7 neisseria m
Q81QR6 bacillus an
Q813E1 pseudomonas
Q8KN7 vibrio chol
Q89M84 bradyrhizob
Q81312 pseudomonas
Q8UB58 agrobacteri
Q8CDK4 lactococcu
Q7VUM9 bordetella
Q88F01 pseudomonas
Q89R44 bradyrhizob
Q7UHK4 rhodopirell
Q7U3U8 synechococc
Q88NR2 pseudomonas
Q8XS32 ralstonia s
Q849M8 arabidopsis
Q8HYT2 pseudomonas
Q8H1R4 thermoplasm
Q82TV9 rhizobium m
Q87V80 pseudomonas
Q89NU4 bradyrhizob
Q83D20 coxiella bu
Q82TW9 rhizobium m
Q8FK62 escherichia
Q8C990 arabidopsis
Q8XCE6 escherichia
Q87ZE5 sulfolobus
Q8C991 arabidopsis
Q811R8 pseudomonas
Q830R8 rhizobium m
Q8Z8R3 salmonella
Q8Z2T6 salmonella
Q874P0 sulfolobus
Q8A7Z3 caulobacter

90 285 9.9 305 16 Q98N82 rhizobium 1
 91 284 9.9 303 16 Q8XU8 ralstonia s
 92 282.5 9.9 329 13 Q7SXU4 Q7sxj4 brachydanio
 93 281.5 9.8 298 16 Q8X8D4 Q8x8d4 escherichia
 94 280 9.8 295 16 Q92PN5 Q92pn5 rhizobium m
 95 279.5 9.8 292 16 Q8Z8Q9 Q8z8q9 salmonella
 96 279.5 9.8 345 5 Q86R98 Q86r98 drosophila
 97 279 9.7 298 16 Q92T07 Q92tg7 rhizobium m
 98 278 9.7 313 16 Q8U8K6 Q8u8k6 agrobacteri
 99 277.5 9.7 359 16 Q8FBG0 Q8fbg0 escherichia
 100 277 9.7 285 16 Q7U3K2 Q7u3k2 synechococc

ALIGNMENTS

RESULT 1

Q98XK2 PRELIMINARY; PRT; 547 AA.
 ID Q98XK2
 AC Q98XK2; (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Cytokine-like nuclear factor n-pac.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA New L., Han J.;
 RT "A novel cytokine-like nuclear factor, n-pac";
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF326966; AAK1524.1; -;
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0004616; F:phosphoglucanate dehydrogenase (decarboxyla. . .; IEA.
 DR GO; GO:0006098; P:pentose-phosphate shunt; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR006115; 6PGD NAD.
 DR InterPro; IPR00637; AT hook.
 DR InterPro; IPR00313; PWMP domain.
 DR Pfam; PF02178; AT hook; 1.
 DR Pfam; PF03446; NAD binding_2; 1.
 DR Pfam; PF00855; PWMP; 1.
 DR SMART; SM00384; AT hook; 1.
 DR PROSITE; PS50812; PWMP; 1.
 DR SEQUENCE 547 AA; 59827 MW; C7D785CCBF83204A CRC64;

Query Match 98.1%; Score 2811; DB 4; Length 547;
 Best Local Similarity 98.7%; Pred. No. 4.9e-189;
 Matches 546; Conservative 0; Mismatches 1; Indels 6; Gaps 1;

QY 1 MAAVSLRLGDLVWGKLGRYPPWPKIVNPPKDLKKPRGKCCFFVKFFGTEDHAWIKVEQL 60
 DB 1 MAAVSLRLGDLVWGKLGRYPPWPKIVNPPKDLKKPRGKCCFFVKFFGTEDHAWIKVEQL 60
 QY 61 KPYHAHEEMIKKNGKRFQOAVDAVEEFLRRAKGDQTS SHNSDDKNRNSSEERSRP 120
 DB 61 KPYHAHEEMIKKNGKRFQOAVDAVEEFLRRAKGDQTS SHNSDDKNRNSSEERSRP 120
 QY 121 NSGDEKRLSLSEGKVKKNMGEKKRVSSGSSERGSKPLKRAQEQSPKRGPRPKDEK 180
 DB 121 NSGDEKRLSLSEGKVKKNMGEKKRVSSGSSERGSKPLKRAQEQSPKRGPRPKDEK 180
 QY 181 LTPESSTVKGMAAGPMAAFKMQPTASEPVKADPHFHHLLSOTKPAVCYQAITKKLK 240
 DB 181 LTPESSTVKGMAAGPMAAFKMQPTASEPVKADPHFHHLLSOTKPAVCYQAITKKLK 240
 QY 241 ICEETGSGTQAADSTAVNGSIPTDKKIGFLGLMGSGIVSNLLKMGHTVTVMNRTA 300
 DB 241 ICEETGSGTQAADSTAVNGSIPTDKKIGFLGLMGSGIVSNLLKMGHTVTVMNRTA 300

QY 301 EKCDLFIQEGARLGTPTAEVSTCDITFACVSDPKAAKDLVLGPGSVLQGIIRPGKCYVDM 360
 DB |||||-----EGARLGTPTAEVSTCDITFACVSDPKAAKDLVLGPGSVLQGIIRPGKCYVDM 354
 QY 361 STVDADTVTELAQVTVSRGGFLEAPVSGNOQLSNDGMLVILAAAGDRLGYEDCSSCFQAM 420
 DB |||||-----STVDADTVTELAQVTVSRGGFLEAPVSGNOQLSNDGMLVILAAAGDRLGYEDCSSCFQAM 414
 QY 421 GKTSPFLGVEGNAAKMMLIVNMVQGSFMATIAEGLTLAHTVQSQOQTLDLILNQGLASI 480
 DB |||||-----GKTSPFLGVEGNAAKMMLIVNMVQGSFMATIAEGLTLAHTVQSQOQTLDLILNQGLASI 474
 QY 481 FLQKQCNILQGNFKPDPFLYKIQKDLRLAIALGDAVNHPTPMAAAANEVYKRAKDQS 540
 DB |||||-----FLQKQCNILQGNFKPDPFLYKIQKDLRLAIALGDAVNHPTPMAAAANEVYKRAKDQS 534
 QY 541 DNDMSAVRAYIH 553
 DB |||||-----DNDMSAVRAYIH 547

RESULT 2

Q922P9 PRELIMINARY; PRT; 546 AA.
 ID Q922P9
 AC Q922P9;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE RIKEN cDNA 393040K13 gene.
 GN NPAC OR 393040K13RIK.
 OS Mus musculus (Mouse)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC006893; AAH06893.1; -;
 DR MGD; MGI:1921272; Npac.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0004616; F:phosphoglucanate dehydrogenase (decarboxyla. . .; IEA.
 DR GO; GO:0006098; P:pentose-phosphate shunt; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR006115; 6PGD NAD.
 DR InterPro; IPR00637; AT hook.
 DR InterPro; IPR00313; PWMP domain.
 DR Pfam; PF02178; AT hook; 1.
 DR Pfam; PF03446; NAD binding_2; 1.
 DR Pfam; PF00855; PWMP; 1.
 DR SMART; SM00384; AT hook; 1.
 DR PROSITE; PS50812; PWMP; 1.
 DR SEQUENCE 546 AA; 59715 MW; F5D2090DE1B64723 CRC64;

Query Match 96.6%; Score 2767.5; DB 11; Length 546;
 Best Local Similarity 97.3%; Pred. No. 5.6e-186;
 Matches 538; Conservative 5; Mismatches 3; Indels 7; Gaps 2;

QY 1 MAAVSLRLGDLVWGKLGRYPPWPKIVNPPKDLKKPRGKCCFFVKFFGTEDHAWIKVEQL 60
 DB 1 MAAVSLRLGDLVWGKLGRYPPWPKIVNPPKDLKKPRGKCCFFVKFFGTEDHAWIKVEQL 60
 QY 61 KPYHAHEEMIKKNGKRFQOAVDAVEEFLRRAKGDQTS SHNSDDKNRNSSEERSRP 120
 DB 61 KPYHAHEEMIKKNGKRFQOAVDAVEEFLRRAKGDQTS SHNSDDKNRNSSEERSRP 120
 QY 121 NSGDEKRLSLSEGKVKKNMGEKKRVSSGSSERGSKPLKRAQEQSPKRGPRPKDEK 180
 DB 121 NSGDEKRLSLSEGKVKKNMGEKKRVSSGSSERGSKPLKRAQEQSPKRGPRPKDEK 179
 QY 181 LTPESSTVKGMAAGPMAAFKMQPTASEPVKADPHFHHLLSOTKPAVCYQAITKKLK 240
 DB |||||-----LTPESSTVKGMAAGPMAAFKMQPTASEPVKADPHFHHLLSOTKPAVCYQAITKKLK 240

Db 180 LTIPESSTVKGMMAGPMAAFKMQPTATEPVKADDPHFHFLLSQTEKPAVCYQAITKKLK 239
 QY 241 ICEBTGSTSIQAADSTAVNGSITPTDKKIGFLGLGMSGIVSNLLKMGHTVTVMNR 300
 Db 240 ICEBTGSTSIQAADSTAVNGSITPTDKKIGFLGLGMSGIVSNLLKMGHTVTVMNR 299
 QY 301 EKCDLPIQEGARLGRTPAEVSTCDITFACVSDPKAAKDLVLGSPGVLGIRPGKYVDM 360
 Db 300 EK-----EGARLGRTPAEVSTCDITFACVSDPKAAKDLVLGSPGVLGIRPGKYVDM 353
 QY 361 STVDADTVELAQVTVSGRFLAPVSGNQLSNDGMLVILAAAGDGLYEDCSCFOAM 420
 Db 354 STVDADTVELAQVTVSGRFLAPVSGNQLSNDGMLVILAAAGDGLYEDCSCFOAM 413
 QY 421 GKTSFFLGEVGNAAKMWLIIVNMVQGSFMATIAEGLTLAHTVTSQOQLDILNQQLASI 480
 Db 414 GKTSFFLGEVGNAAKMWLIIVNMVQGSFMATIAEGLTLAHTVTSQOQLDILNQQLASI 473
 QY 481 FLDQKQCNILQGNFKPDPFYLYIKQDLRLAIALGDAVNHPTPMAAAANEVYKRAKALDQS 540
 Db 474 FLDQKQCNILQGNFKPDPFYLYIKQDLRLAIALGDAVNHPTPMAAAANEVYKRAKALDQS 533
 QY 541 DNDMSAVRYAYIH 553
 Db 534 DNDMSAVRYAYIH 546

RESULT 3
 Q9CYQ1
 ID Q9CYQ1 PRELIMINARY; PRT; 546 AA.
 AC Q9CYQ1
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE 3930401K13rik protein.
 GN NPAC OR 3930401K13RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CS7BL/6J; TISSUE=Embryo;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staebli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombauts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection";
 RL Nature 409:685-690(2001).
 DR EMBL: AK014456; BAB29363.1; --
 DR MGD: MGI:1921272; Npac.
 DR GO: GO:0005634; C:nucleus; IEA.
 DR GO: GO:0003677; F:DNA binding; IEA.
 DR GO: GO:0004616; P:phosphogluconate dehydrogenase (decarboxyla. . .); IEA.
 DR GO: GO:0006098; P:pentose-phosphate shunt; IEA.
 DR GO: GO:0006355; P:regulation of transcription; IEA, DNA-dependent; IEA.
 DR InterPro: IPR006115; 6PGD_NAD.
 DR InterPro: IPR000637; AT_hook.

DR InterPro: IPR000313; PWWP domain.
 DR Pfam: PF02178; AT_hook; 1.
 DR Pfam: PF03446; NAD_binding_2; 1.
 DR Pfam: PF00855; PWWP; 1.
 DR SMART: SM00384; AT_hook; 1.
 DR PROSITE: PS00812; PWWP; 1.
 SQ SEQUENCE 546 AA; 59744 MW; B3AC1562477ABC1F CRC64;

Query Match 95.8%; Score 2745.5; DB 11; Length 546;
 Best Local Similarity 96.6%; Pred. No. 2e-184;
 Matches 534; Conservative 8; Mismatches 4; Indels 7; Gaps 2;

QY 1 MAAVSLRLGDLVWVKLGRLYPWPFGKI VNPDPKDKKPRGKKCFVKKFFGTEGDAWTKVEQL 60
 Db 1 MAAVSLRLGDLVWVKLGRLYPWPFGKI VNPDPKDKKPRGKKCFVKKFFGTEGDAWTKVEQL 60
 QY 61 KPYHAHEEMIKKNGKRFQQA VDAVEEFLRAKGDQTSNHSDDDKNRNRSBERSRP 120
 Db 61 KPYHAHEEMIKKNGKRFQQA VDAVEEFLRAKGDQTSNHSDDDKNRNRSBERSRP 120
 QY 121 NSGDEKRLSLSEGKVKVMGEGKKRVSSGSSERGSKPLKRAQOSPRKGRPPKDEKD 180
 Db 121 NSGDEKRLSLSEGKVKVMGEGKKRVSSGSSERGSKPLKRAQOSPRKGRPPKDEKD 179
 QY 181 LTIPESSTVKGMMAGPMAAFKMQPTASBPVKDADPHEHFLLSQTEKPAVCYQAITKKLK 240
 Db 180 LTIPESSTVKGMMAGPMAAFKMQPTATEPVKADDPHFHFLLSQTEKPAVCYQAITKKLK 239
 QY 241 ICEBTGSTSIQAADSTAVNGSITPTDKKIGFLGLGMSGIVSNLLKMGHTVTVMNR 300
 Db 240 ICEBTGSTSIQAADSTAVNGSITPTDKKIGFLGLGMSGIVSNLLKMGHTVTVMNR 299
 QY 301 EKCDLPIQEGARLGRTPAEVSTCDITFACVSDPKAAKDLVLGSPGVLGIRPGKYVDM 360
 Db 300 EK-----EGARLGRTPAEVSTCDITFACVSDPKAAKDLVLGSPGVLGIRPGKYVDM 353
 QY 361 STVDADTVELAQVTVSGRFLAPVSGNQLSNDGMLVILAAAGDGLYEDCSCFOAM 420
 Db 354 STVDADTVELAQVTVSGRFLAPVSGNQLSNDGMLVILAAAGDGLYEDCSCFOAM 413
 QY 421 GKTSFFLGEVGNAAKMWLIIVNMVQGSFMATIAEGLTLAHTVTSQOQLDILNQQLASI 480
 Db 414 GKTSFFLGEVGNAAKMWLIIVNMVQGSFMATIAEGLTLAHTVTSQOQLDILNQQLASI 473
 QY 481 FLDQKQCNILQGNFKPDPFYLYIKQDLRLAIALGDAVNHPTPMAAAANEVYKRAKALDQS 540
 Db 474 FLDQKQCNILQGNFKPDPFYLYIKQDLRLAIALGDAVNHPTPMAAAANEVYKRAKALDQS 533
 QY 541 DNDMSAVRYAYIH 553
 Db 534 DNDMSAVRYAYIH 546

RESULT 4
 Q9BT11
 ID Q9BT11 PRELIMINARY; PRT; 523 AA.
 AC Q9BT11
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Similar to RIKEN cDNA 3930401K13 gene (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RA Strausberg R.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC003693; AAH03693.1; --
 DR GO: GO:0005634; C:nucleus; IEA.
 DR GO: GO:0003677; F:DNA binding; IEA.

DR GO; GO:0004616; F:phosphogluconate dehydrogenase (decarboxyla. . .; IEA.
 DR GO; GO:0006099; P:pentose-phosphate shunt; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR006115; 6PGD_NAD.
 DR InterPro; IPR000637; AT hook.
 DR InterPro; IPR000313; PWP domain.
 DR Pfam; PF02178; AT hook; 1.
 DR Pfam; PF03446; NAD_binding_2; 1.
 DR Pfam; PF00855; PWP; 1.
 DR SMART; SM00384; AT hook; 1.
 DR PROSITE; PS00812; PWP; 1.
 FT NON_TER
 SQ SEQUENCE 523 AA; 57215 MW; ED3E2853938A211F CRC64;
 Query Match 93.6%; Score 2693.5; DB 4; Length 523;
 Best Local Similarity 96.7%; Pred. No. 4.2e-180;
 Matches 522; Conservative 0; Mismatches 1; Indels 17; Gaps 1;
 QY 14 GKLGRYPWPVGKIVNPPDKLKPGRKKCFVFFGTEHAWIKVEOLKPYHAKEEMIKI 73
 Db 1 GKLGRYPWPVGKIVNPPDKLKPGRKKCFVFFGTEHAWIKVEOLKPYHAKEEMIKI 60
 QY 74 NKGKRFQQAQVDAVEEFLRAKGDQTSNHSDDKNRNSSEERSPNSGDEKRLSLSE 133
 Db 61 NKGKRFQQAQVDAVEEFLRAKGDQTSNHSDDKNRNSSEERSPNSGDEKRLSLSE 120
 QY 134 GKVKQMGKGRVSSGSSGSKPLKRAQSGPKRGRPPKDEKDLTIPSSVTYKGM 193
 Db 121 GKVKQMGKGRVSSGSSGSKPLKRAQSGPKRGRPPKDEKDLTIPSSVTYKGM 180
 QY 194 AGPMAAFKQPTASEPVKADPHFHLFSQTEKPAVCQAATKLCICEETGTSIOA 253
 Db 181 AGPMAAFKQPTASEPVKADPHFHLFSQTEK-----ETGTSIOA 223
 QY 254 ADSTAVNGSITPTDKKIGFLGLGMSGIVSNLLKMGHTVTVMNRTAEKCDLFIQSGARL 313
 Db 224 ADSTAVNGSITPTDKKIGFLGLGMSGIVSNLLKMGHTVTVMNRTAEKCDLFIQSGARL 283
 QY 314 GRTPAEVSCTDITFACVSDPKAAKDLVLPSPVLOGIRPKCYVDMSTVDATVTELAQ 373
 Db 284 GRTPAEVSCTDITFACVSDPKAAKDLVLPSPVLOGIRPKCYVDMSTVDATVTELAQ 343
 QY 374 VIVSRGGRFLFEPVSGNQSLNDGMLVLAAGDGLYEDCSCFQAMGKTSFFLGEVGN 433
 Db 344 VIVSRGGRFLFEPVSGNQSLNDGMLVLAAGDGLYEDCSCFQAMGKTSFFLGEVGN 403
 QY 434 AKMMLIVNVVQGSFMATIAEGLTLAHTVGTQSQTLLDILNQGLASIFLDKQCNILQGN 493
 Db 404 AKMMLIVNVVQGSFMATIAEGLTLAHTVGTQSQTLLDILNQGLASIFLDKQCNILQGN 463
 QY 494 FKPDFVLKYIQDLRLAIALGDVNHPTPMAAAANEVYKRAKALDOSDNDMSAVYRAYTH 553
 Db 464 FKPDFVLKYIQDLRLAIALGDVNHPTPMAAAANEVYKRAKALDOSDNDMSAVYRAYTH 523
 RESULT 5
 Q8T079 PRELIMINARY; PRT; 602 AA.
 AC Q8T079; Q9VL51;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE LD22344p (CG4747-PA).
 GN BEST:LD22483 OR BEST:LD29743 OR CG4747.
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,

Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 Nunoo J., Pacle J., Paragas V., Park S., Phouanavong S., Wan K.,
 Yu C., Lewis S.E., Rubin G.M., Celniker S.;
 Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 Brannon R.C., Rogers Y.H., Blazey J.H., Champe C., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Baldwin D.,
 Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Gabor G.L.,
 Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 Beeson K.Y., Botchan M.R., Boush B.P., Brokstein P., Brotter P.,
 Borkova D., Bontas D.A., Butler H., Cadieu E., Center A., Chandra I.,
 Burtis K.C., Busam D.A., Butler C., Davenport L.B., Davies P.,
 Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Dietz S.M.,
 de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
 Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,
 Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 Williams S.M., Woodger, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 "The genome sequence of Drosophila melanogaster";
 Science 287:2185-2195(2000).
 RN [3]
 RN SEQUENCE FROM N.A.
 RP Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 Hradscky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
 Tupy J.L., Bergman C.M., Berman B.P., Carlson J.W., Celniker S.E.,
 Clamp M.E., Drysdale R.A., Emmert D., Frise E., de Grey A.D.N.J.,
 Harris N.L., Kronmiller B., Marshall B., Millburn G.H., Richter J.,
 Russo S., Searle S.M.J., Smith E., Shu S., Smutniak F.,
 Whitfield E.J., Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J.,
 Lewis S.E.;
 "Annotation of Drosophila melanogaster genome";
 Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RN SEQUENCE FROM N.A.
 RA FlyBase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RN SEQUENCE FROM N.A.
 RA FlyBase;
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY069497; AAL39642.1; -;
 DR EMBL; AE003627; AAF52846.3; -;
 DR FlyBase; FBgn0043456; BEST:LD22483.
 DR GO; GO:0004616; F:phosphogluconate dehydrogenase (decarboxyla. .; IEA.
 DR GO; GO:0006098; P:pentose-phosphate shunt; IEA.
 DR InterPro; IPR006115; 6PGD_NAD.
 DR InterPro; IPR000313; PWP domain.

```
DR PFam; PF03446; NAD_binding_2; 1.
DR PFam; PF00855; PWWP; 1.
DR SMART; SM00293; PWWP; 1.
DR PROSITE; PS00812; PWWP; 1.
SQ SEQUENCE 602 AA; 65452 MW; E004EE610C23625 CRC64;

Query Match      27.8%; Score 796.5; DB 5; Length 602;
Best Local Similarity 32.8%; Pred. No. 1.8e-47;
Matches 192; Conservative 110; Mismatches 225; Indels 59; Gaps 13;

QY 10 DLVWGLKGRYPWPGKIVNPKDL---KKPKGKKCFVKVFGTEHDAMKVEQKPYHA- 65
Db 24 DLIWAKMGFTWPGMIVDPDLILSQORRANTKCVF--PFGSRNFAWIEENNKKPFEP 81
QY 66 HKEEMIKINKKRFQQAQVDAVEEFLRRAKGKDQTSNHSDDKNRNSSEERSPNSGDE 125
Db 82 WKEELAKVSKPAAFRHMTDIEKYIDDPAEVDEQVNSCGAPNHAATEADFKIR--DGLD 139
QY 126 KRKLSLGKGVKYNMGEGKKRVSSGSSRG-----SKSPL-----KRAQRQS 167
Db 140 SEEVIGBEATADGNGGVVAVHV--GSPDEGDGLDVEINADSSASPVTSPTVTKAAGKRT 197
QY 168 PR-----KRGPRPKDEKDLITPESSTVKGMMAGPMAAPKWOPTAS-BPVKQADP 215
Db 198 PKAKSVAATSVKSTKGSAKSAQKRTSAQQSP-----SGPSNAKRGRDVSGEALQDADE 252
QY 216 -----HFHFLLSQTEKPAVCYQALTKKLKICEETGTSIQAADSTAVNGSIT 264
Db 253 ASSTPTGRRRVETDALLASIAAKAPNAIALDRPVTRPQAVIDMSRSNTLIADRIDV 312
QY 265 PTDKKIGFLGLGMSGIVSNLLKMGHTVTVNRTAEKCDLFIQEGARLGRTPAEVSTC 324
Db 313 PSEQTGFLGLGMSGIVSNLLKMGHTVTVNRTAEKCDLFIQEGARLGRTPAEVSTC 372
QY 325 DITFACVSDPKAAKDLVLPSPGVQGRIP--GKCYVDMSTVDADTVTELAQVIVSRGRF 382
Db 373 DVIFCCVSDPKGAKDLVFGNGVQLQ-LKDLNKAIVEMSTTIDPDTDIDGEGIKQCGRY 431
QY 383 LEAPVSGNQLSNDGMLVILAAGDRGLYEDCSSCFQAMGKTSFFLGEVGNAAKMLVNM 442
Db 432 LEAQHSRQEAAGMLIILAGGRSVFECHSCFTAKNTFFLGNIGNACKNLLIQT 491
QY 443 VQGSFMATIAEGLTIAHVTSQSOQTLLDILNQGLASIFLDQKQNLQGNFKPDFYLKY 502
Db 492 ILGVSILVGLAEALADRFSLINDIIDIFDLTSMKSPMLLAKGKEMAKGDFNQQLPISH 551
QY 503 IQKDLRLAIALGDVAVNHTPMAAAANEVYKRAKALDOSDNDMSAVY 548
Db 552 MQDRLRLVLAENLDQSMPTVTSITNEVFKTRKLGYSHEHSSAVF 597

RESULT 6
Q84VC8 PRELIMINARY; PRT; 293 AA.
AC Q84VC8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Gamma hydroxybutyrate dehydrogenase-like protein (Hypothetical protein).
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Cooper B., Hutchison D., Park S., Guimil S., Luginbuhl P., Ellero C.,
RA Goff S., Glazebrook J.,
RT "Identification of Rice (Oryza sativa) Proteins Linked to the Cyclin-
RT Mediated Regulation of the Cell Cycle."
RT Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
```

```

[2]
RN RP SEQUENCE FROM N.A.
RC TISSUE-Panicle;
RA Moriguchi K., Ito Y., Yamazaki Y., Kurata N.;
RT "Finding of various plant nuclear proteins using yeast nuclear
RT transportation trap system - a proteomal approach."
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY224558; AA072678.1; -.
DR EMBL; AB110167; BAC78559.1; -.
DR GO; GO:0008442; F:3-hydroxyisobutyrate dehydrogenase activity; IEA.
DR GO; GO:0004616; F:phosphogluconate dehydrogenase (decarboxyla. . .; IEA.
DR GO; GO:0006098; P:pentose-phosphate shunt; IEA.
DR GO; GO:0006573; P:valine metabolism; IEA.
DR InterPro; IPR002204; 3hydroxoisobut_dh.
DR InterPro; IPR006183; 6PGD.
DR InterPro; IPR006115; 6PGD_NAD.
DR Pfam; PF03446; NAD_binding_2; 1.
DR PRINTS; PR00076; 6PGDHRGNASE.
DR PROSITE; PS00895; 3 HYDROXYISOBUT_DH; 1.
KW Hypothetical protein.
SQ SEQUENCE 293 AA; 30496 MW; 2362B7AC4EE52038 CRC64;

Query Match      25.4%; Score 727; DB 10; Length 293;
Best Local Similarity 52.1%; Pred. No. 5e-43;
Matches 146; Conservative 49; Mismatches 85; Indels 0; Gaps 0;

QY 269 KIGFLGLGMSGIVSNLLKMGHTVTVNRTAEKCDLFIQEGARLGRTPAEVSTCITP 328
Db 2 EVGFLGIGMGKAAANLLRHGFRVTVNRTLSKCQELVALGAAGVETPAAVAKCRYTI 61
QY 329 ACVSDPKAAKDLVLPSPGVQGRIPGKCYVDMSTVDADTVTELAQVIVSRGRFLEAPVS 388
Db 62 AMLSDEPSAALSVDPDKGVLEIQEGEGKYVDMSTVDAATSKISEAIKQKGAFVEAPVS 121
QY 389 GNQQLSNDGMLVILAAGDRGLYEDCSSCFQAMGKTSFFLGEVGNAAKMLVNMVQGSFM 448
Db 122 GSKPAEDGQQLVILAAGDKVLYDDMPAFDVLGKKSFFLGEIGNGAKMKLVNMIMSGMM 181
QY 449 ATIAEGLTIAHVTSQSOQTLLDILNQGLASIFLDQKQNLQGNFKPDFYLKYIKDLR 508
Db 182 NALSEGSLADNSGLSPQTLTLDVLDLGAIANPMFKLGPSMLQSGSYNPAEPFLKQKQDMR 241
QY 509 LAIALGDVAVNHTPMAAAANEVYKRAKALDOSDNDMSAVY 548
Db 242 LALALGDENAVSMPTVAASNEAFKARSGLGLGLDLDFAVY 281

RESULT 7
Q94B07 PRELIMINARY; PRT; 289 AA.
AC Q94B07;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Gamma hydroxybutyrate dehydrogenase (EC 1.1.1.61).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC Breithreuz K.E., Van Cauwenberghe O.R., Allen W.L., Jakobs C.,
RA Talibi D., McLean M.D., Andre B., Shelp B.J.;
RT "Plant gamma-hydroxybutyrate dehydrogenase, a novel fermentation
RT enzyme."
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY041483; AAQ94781.1; -.
DR GO; GO:0008442; F:3-hydroxyisobutyrate dehydrogenase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0004616; F:phosphogluconate dehydrogenase (decarboxyla. . .; IEA.
DR GO; GO:0006098; P:pentose-phosphate shunt; IEA.
DR GO; GO:0006573; P:valine metabolism; IEA.
```


DR EMBL; AY085690; AAMG2909.1; --
 DR EMBL; BT008734; AAP42747.1; --
 DR GO; GO:0004616; P:phosphogluconate dehydrogenase (decarboxyla. . .; IEA.
 DR GO; GO:0006038; P:pentose-phosphate shunt; IEA.
 DR InterPro; IPR006115; 6PGD_NAD.
 DR Pfam; PF03446; NAD_binding_2; 1.
 DR Hypothetical protein.
 SQ SEQUENCE 343 AA; 36287 MW; D012C16284C5FA52 CRC64;
 Query Match 23.3%; Score 667; DB 10; Length 343;
 Best Local Similarity 39.4%; Pred. No. 1e-38;
 Matches 142; Conservative 64; Mismatches 116; Indels 38; Gaps 4;
 QY 196 PMAAFKWOPTASEPVKADPHHFLLSQTEKPAVC-----FL-----SKPQICLAYRVYSSLOSTPTSTRDELGTVS 250
 Db 8 PRIPRFRPKPISP-----FL-----SKPQICLAYRVYSSLOSTPTSTRDELGTVS 53
 QY 251 IQAADSTAVNGSITPTDKKIGFLGLGMSGIVSNLLKMGHTVTVNRTAKCDLFIQEG 310
 Db 54 -----IGFLGMSGVPMQAQNLKAGCDVTWNRKSKCDPLVGLG 94
 QY 311 ARLGRTPAEVSTCDITFACVSDPKAAKDLVLGSGVLQIRPKCYVDMSTVDADTYTE 370
 Db 95 AKYKSSPEVATCDLTFTAMLADPESADVACGRNGAIFGSSGKGYVDVSTVDVASSIL 154
 QY 371 LAQIVSRGGRFLAPVSGNQQLSNDGMLAVILAAGDRGLYEDCSCFOAMGKTFFLGEV 430
 Db 155 ISKQIKOTGALFLAPVSGSKKPAEDGQILFTAGDKPLYEKAAPFLDIMGSKFYLGEV 214
 QY 431 GNAAMMLIVNMGVSGFMATIAEGLTLTAHVTSQSQTLDDILNQGLASIFLDQKCNIL 490
 Db 215 GNGAAMKLIVNMGVSGFMATIAEGLTLTAHVTSQSQTLDDILNQGLASIFLDQKCNIL 490
 QY 491 QGNFKPDLKAYIOKDLRLATLADGAVNHPPTMAAANEVYKRAKALDOSNDMSAVYRA 550
 Db 275 KSVYTPAPPLKHOQDMRLAELAESVSQSTPIAAANELYKVAKSYGLSDEDFSAVIEA 334

RESULT 10

Q94A74 Q94A74 PRELIMINARY; PRT; 387 AA.
 AC Q94A74
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE F113.35.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Khan S., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Kim C.,
 RA Shinn P., Altafi H., Bei B., Chin C., Chiou J., Choi E., Conn L.,
 RA Conway A., Gonzales A., Hansen B., Howng B., Koo T., Lam B., Lee J.,
 RA Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N., Nguyen M.,
 RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
 RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
 RA Ecker J.R.;
 RT "Genomic sequence for Arabidopsis thaliana BAC F113 from chromosome
 I.1";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.

RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
 RA Khan S., Kim C., Altafi H., Bei B., Chin C., Chiou J., Choi E.,
 RA Conn L., Conway A., Gonzales A., Hansen B., Howng B., Koo T., Lam B.,
 RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N.,
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
 RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
 RA Theologis A., Ecker J.;
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC022492; RAF79475.1; --
 DR GO; GO:0004616; P:phosphogluconate dehydrogenase (decarboxyla. . .; IEA.
 DR GO; GO:0006038; P:pentose-phosphate shunt; IEA.
 DR InterPro; IPR006115; 6PGD_NAD.
 DR Pfam; PF03446; NAD_binding_2; 1.
 SQ SEQUENCE 387 AA; 41484 MW; 0445432C58CA0137 CRC64;
 Query Match 22.5%; Score 644; DB 10; Length 387;
 Best Local Similarity 36.3%; Pred. No. 5e-37;
 Matches 146; Conservative 65; Mismatches 121; Indels 70; Gaps 6;
 QY 186 SSTVKGM---MAGPMAAFKWOPTASEPVKADPHHFLLSQTEKPAVC-----YQAITK 237
 Db 10 SSSKAMALCICRIPRFRPKPISP-----FL-----SKPQICLAYRVYSSLSQS 55
 QY 238 KLKICEERTGSTSIQADSTAVNGSITPTDKKIGFLGLGMSGIVSNLLKMGHTVTVN 297
 Db 56 TTPSTRDELGTVS-----IGFLGMSGVPMQAQNLKAGCDVTWVN 96
 QY 298 RTAEKCDLFIQEGARLGRTPAEVSTCDITFACVSDPKAAK----- 338
 Db 97 RTKSKCDPLVGLGAKYKSSPEVATCDLTFTAMLADPESAVRMGHFCLRLFFPFHSLK 156
 QY 339 -----DLVLGSGVLQIRPKCYVDMSTVDADTYTELAQVIVSRGRFLAPVS 388
 Db 157 QKLVAFQCIDVACGKNGAIFGSSGKGYVDVSTVDVASSILSKQIKDGTGALFLAPVS 216
 QY 389 GNOQLSNDGMLAVILAAGDRGLYEDCSCFOAMGKTFFLGEVGNAAKMLIVNMGVSGFM 448
 Db 217 GSKKPAEDGQILFTAGDKPLYEKAAPFLDIMGSKFYLGEVGNAAKMLIVNMGVSGFM 276
 QY 449 ATIAEGLTLTAHVTSQSQTLDDILNQGLASIFLDQKCNILQGNFKPDLFYKIQKDLR 508
 Db 277 ASFAEGLLSQKVGLDPNVLVEVVSQAINAPMTSLKSPMSIKSVYTPAPPLKHOQDMR 336
 QY 509 LAIALGDAVNHPPTMAAANEVYKRAKALDOSNDMSAVYRA 550
 Db 337 LAIGLAESVSQSTPIAAANELYKVAKSYGLSDEDFSAVIEA 378
 RESULT 11
 Q94A74 Q94A74 PRELIMINARY; PRT; 248 AA.
 AC Q94A74
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE AT3g25530/MWL2.15.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kim C.J., Chen H., Cheuk R., Koesema E., Meyers M.C., Banh J.,
 RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
 RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
 RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 RT "Arabidopsis cDNA clones";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

```

DR EMBL; AY049298; AAK83640.1; -.
DR GO; GO:0008442; F:3-hydroxyisobutyrate dehydrogenase activity; IEA.
DR GO; GO:0004616; F:phosphogluconate dehydrogenase (decarboxyla. .; IEA.
DR GO; GO:0006098; P:pentose-phosphate shunt; IEA.
DR GO; GO:0006573; P:valine metabolism; IEA.
DR InterPro; IPR002204; 3hydroxisobut_dh.
DR InterPro; IPR006183; 6PGD.
DR InterPro; IPR006115; 6PGD_NAD.
DR Pfam; PF03446; NAD_binding_2; 1.
DR PRINTS; PR00076; 6PGDHDRGNASE.
DR PROSITE; PS00895; 3_HYDROXYISOBUT_DH; 1.
SQ SEQUENCE 248 AA; 26336 MW; C96912B3E2A7E677 CRC64;

Query Match 19.6%; Score 562; DB 10; Length 248;
Best Local Similarity 51.7%; Pred. No. 1.5e-31;
Matches 109; Conservative 36; Mismatches 66; Indels 0; Gaps 0;

QY 269 KIGFLGLGMSGIVSNLLKMGHTVTVNRTAEKCDLFTQEGARLGRTPAEVVSCTDITF 328
DB 2 EVGFLGLGMSGIVSNLLKMGHTVTVNRTAEKCDLFTQEGARLGRTPAEVVSCTDITF 61
QY 329 ACVSDPKAAKDLVLGSPGVLOGIRPGKCYVDMSTVDADTVTBLAQVIVSRGGRFLEAPVS 388
DB 62 AMLSDPCAALSVVDFKGVLEQICGKGYIDMSTVDAETSLKINEAITGKGRFVEGPVS 121
QY 389 GNQQLSNDGMLVILAAGDRGLYEDCSSCFQAMGKTSFELGVEGNAKMLIVNMVQGSFM 448
DB 122 GSKKPAEDQQLIILAAGDKALFEESIPAFDVLGKRSFYLGQVNGAKMLIVNMVQGSFM 181
QY 449 ARTAELTLAHVTGSGSQOQLLILNQGLAS 479
DB 182 NAFSEGLVLADKSLSSDFTLLDLDLGATN 212

RESULT 12
Q8LQJ7 PRELIMINARY; PRT; 343 AA.
ID Q8LQJ7
AC Q8LQJ7;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-JUN-2003 (TREMBLrel. 22, Last sequence update)
DE Putative gamma hydroxybutyrate dehydrogenase.
GN B1151A10.17.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1, BAC
clone:B1151A10."
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF003413; BAB91904.1; -.
DR Gramene; Q8LQJ7; -.
DR GO; GO:0008442; F:3-hydroxyisobutyrate dehydrogenase activity; IEA.
DR GO; GO:0004616; F:phosphogluconate dehydrogenase (decarboxyla. .; IEA.
DR GO; GO:0006098; P:pentose-phosphate shunt; IEA.
DR GO; GO:0006573; P:valine metabolism; IEA.
DR InterPro; IPR002204; 3hydroxisobut_dh.
DR InterPro; IPR006115; 6PGD_NAD.
DR Pfam; PF03446; NAD_binding_2; 1.
DR PROSITE; PS00895; 3_HYDROXYISOBUT_DH; 1.
SQ SEQUENCE 343 AA; 35481 MW; 195ABBEDA048F84D CRC64;

Query Match 19.5%; Score 560; DB 10; Length 343;
Best Local Similarity 40.5%; Pred. No. 3.3e-31;
Matches 125; Conservative 49; Mismatches 83; Indels 52; Gaps 4;

QY 269 KIGFLGLGMSGIVSNLLKMGHTVTVNRTAEKCDLFTQEGARLGRTPAEVVSCTDITF 328
DB 2 EVGFLGLGMSGIVSNLLKMGHTVTVNRTAEKCDLFTQEGARLGRTPAEVVSCTDITF 61
QY 329 ACVSDPKAAKDLVLGSPGVLOGIRPGKCYVDMSTVDADTVTBLAQVIVSRGGRFLEAPVS 388
DB 62 AMLSDPCAALSVVDFKGVLEQICGKGYIDMSTVDAETSLKINEAITGKGRFVEGPVS 121
QY 389 GNQQLSNDGMLVILAAGDRGLYEDCSSCFQAMGKTSFELGVEGNAKMLIVNMVQGSFM 448
DB 122 GSKKPAEDQQLIILAAGDKALFEESIPAFDVLGKRSFYLGQVNGAKMLIVNMVQGSFM 181
QY 449 ARTAELTLAHVTGSGSQOQLLILNQGLAS 479
DB 182 NAFSEGLVLADKSLSSDFTLLDLDLGATN 212

```

```

DB 52 KVGFGLGIMGAPMASNLINAGCDVTVNVRTSKCDPLLSLCAKYEPSPADVASSCDVTF 111
QY 329 ACVSDPKAAKDLVLGSPGVLOGIRPGKCYVDMSTVDADTVTBLAQVIVSRGGRFLEAPVS 388
DB 112 AMLADPESAVEVACGANGAAGMAPGKGYVDVSTVDAATSKLIGHITSTGASFLERC-- 169
QY 389 GNQQLSNDGMLVILAAGDRGLYEDCSSCFQAMGKTSFELGVEGNAKMLIVNMVQGSFM 424
DB 170 -----DFSSCTVGPMLTRYLFSLRQFQAQKQQKMGCSSTLP 206
QY 425 ---PFLGVEGNAKMLIVNMVQGSFMATIAEGLTLAHVTGSGSQOQLLILNQGLASIF 481
DB 207 QSRFFLGVDGKADMKLVNMVQGSFMVFSSEGLLSLSEKVGLDPTNLVEVISQGAISAPM 266
QY 482 LDQKCNILQGNFKPDEVLKYIQDLRLAIALGDAVNHPTPMAAANEVYKRAKALDOSD 541
DB 267 FSLGSPVMWKAAYPTAFELKHOQDLRLALALAESVSQSIPTVAANELYKVAKSLGLAD 326
QY 542 NDMSAVYRA 550
DB 327 QDFSAVIEA 335

RESULT 13
Q8TT25 PRELIMINARY; PRT; 300 AA.
ID Q8TT25
AC Q8TT25;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE 3-hydroxyisobutyrate dehydrogenase.
GN MA0614.
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA FitzHugh W., Calvo S., Engels R., Smirnov S., Ancoor D., Brown A.,
RA Allen N., Naylor J., Stange-Thomann N., DeArellano K., Johnson R.,
RA Linton L., McSwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
RA Hederich R., Ingram-Smith C., Kuettnner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
and physiological diversity."
RL Genome Res. 12:532-542(2002).
DR EMBL; AE010722; AAM04058.1; -.
DR GO; GO:0004616; F:phosphogluconate dehydrogenase (decarboxyla. .; IEA.
DR GO; GO:0006098; P:pentose-phosphate shunt; IEA.
DR InterPro; IPR006183; 6PGD.
DR InterPro; IPR006115; 6PGD_NAD.
DR Pfam; PF03446; NAD_binding_2; 1.
DR PRINTS; PR00076; 6PGDHDRGNASE.
DR Complete proteome.
SQ SEQUENCE 300 AA; 32287 MW; FB1E7A8A8A8A7E CRC64;

Query Match 13.9%; Score 397; DB 17; Length 300;
Best Local Similarity 31.8%; Pred. No. 7.8e-20;
Matches 90; Conservative 47; Mismatches 144; Indels 2; Gaps 1;

QY 267 DKKI--GFLGLGMSGIVSNLLKMGHTVTVNRTAEKCDLFTQEGARLGRTPAEVVSCT 324
DB 7 DKNITVGVLGMSGFSFASNLISRGYNVHVYNTKEKAQPLIERGATPHSTPRELASVA 66
QY 325 DITFACVSDPKAAKDLVLGSPGVLOGIRPGKCYVDMSTVDADTVTBLAQVIVSRGGRFLE 384

```



```

Db 67 DIIMTSLTDEAAVNSVAFGEDGLNGAKKGLWLDLSTIDPSSSVKHAABAAKAGLERLD 126
QY 385 APVSGNQQLSNDGMLVILAAAGDRGLYEDSCSCFQAMGKTSFFLGEVGNAAKMWLVNVMQV 444
Db 127 TPVSGSKLASKEGELIILVGGQEVLRKHEKFLNKLGRSVIYLGADGNGHKMKLAINLHL 186
QY 445 GSFMTAIEAGLFLAHTVGTOSQOTLLDILNQGLASIFLDQKQNLQGNFKPDFYLYKIYQ 504
Db 187 GLLAESFSEALVFSQKLGFDKTFVETINNTPIRNVISQKGRIVEGNEFPAFSLNLIA 246
QY 505 KDLRLAIALGDVNHPTPMAAANEVYKRAKALDQSDNDMSAV 547
Db 247 KDLRLVNEQITKTGAILPMTKVSIBEYSRTVQNGEQKDFSVI 289

RESULT 14
O66454 PRELIMINARY; PRT; 288 AA.
AC O66454;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE 3-hydroxyisobutyrate dehydrogenase.
GN HIBD OR AQ_038.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warten P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Snead M.A., Keller M., Aujay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus.";
RL Nature 392:353-358(1998).
DR EMBL; AE000670; AAC06408.1; -.
DR PIR; C70303; C70303.
DR GO; GO:0004616; F:phosphogluconate dehydrogenase (decarboxyla. . .; IEA.
DR GO; GO:0006038; P:pentose-phosphate shunt; IEA.
DR InterPro; IPR008927; 6GDH_C_like.
DR InterPro; IPR006183; 6PGD.
DR Pfam; PF03446; NAD binding 2; 1.
DR PRINTS; PR00076; 6PGDHRGNASE.
DR Complete proteome.
KW SEQUENCE 288 AA; 31788 MW; 60E79AF4954D4294 CRC64;

Query Match 13.7%; Score 392.5; DB 16; Length 288;
Best Local Similarity 31.7%; Pred. No. 1.5e-19;
Matches 90; Conservative 61; Mismatches 132; Indels 1; Gaps 1;

QY 269 KIGFLGLMGSGIVSNLLKMGHTVTVNNTAEKCDLFIQF-GARLGRTPAEVYSTCDIT 327
Db 2 KVGFIGLHGLRAELAKELIEQGVLELVNRTLSKAHEFAKETGAETVSSPADLINKVDV 61
QY 328 FACVSDPKAAKDLVLGSPGVLGIRPGKCYVDMSTVDADTVTELAQVIVSRGRFLAPV 387
Db 62 FVIVFDSQAGEEVIFGEGKLVGDIKGTVDIMTTHYLAQKAYEBELKXGAFYLDAPV 121
QY 388 SGNQQLSNDGMLVILAAAGDRGLYEDSCSCFQAMGKTSFFLGEVGNAAKMWLVNVMQGSF 447
Db 122 LGSVIPALKGELTIVVGDKKEFENKPEKFCRAIYILGEAGSGKMKLVNINVILGGI 181
QY 448 MATIAEGLTIAHVTGQSQOTLLDILNQGLASIFLDQKQNLQGNFKPDFYLYKIQK 507
Db 182 MEVLAIAIGAERAGIDKETLINVLDNGAGKSYILDYKKKLUEDSFTHFSVNLIIYKDL 241
QY 508 RLALGDAVNHPTPMAAANEVYKRAKALDQSDNDMSAVTRAY 551
Db 242 HYAODLIKDLGLFSFTTAAVVKETYGLARKGEGNLDPSAVYKLF 285

```

```

RESULT 15
Q97XZ7 PRELIMINARY; PRT; 289 AA.
ID Q97XZ7;
AC Q97XZ7;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Oxidoreductase.
GN SSO1560.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA She O., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Awayez M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Etrauso G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR EMBL; AE006770; AAK41774.1; -.
DR PIR; G90314; G90314.
DR GO; GO:0008442; F:3-hydroxyisobutyrate dehydrogenase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0004616; F:phosphogluconate dehydrogenase (decarboxyla. . .; IEA.
DR GO; GO:0006038; P:pentose-phosphate shunt; IEA.
DR GO; GO:0006573; P:valine metabolism; IEA.
DR InterPro; IPR002204; 3hydroxisobut_dh.
DR InterPro; IPR006183; 6PGD.
DR InterPro; IPR006115; 6PGD_NAD.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF03446; NAD binding 2; 1.
DR PRINTS; PR00076; 6PGDHRGNASE.
DR PROSITE; PS00895; 3 HYDROXYISOBUT_DH; 1.
DR PROSITE; PS00030; REM_RNP_1; 1.
KW Complete proteome.
SQ SEQUENCE 289 AA; 31025 MW; 4EEA33390300E7C7 CRC64;

Query Match 13.7%; Score 392; DB 17; Length 289;
Best Local Similarity 29.3%; Pred. No. 1.7e-19;
Matches 83; Conservative 68; Mismatches 132; Indels 0; Gaps 0;

QY 269 KIGFLGLMGSGIVSNLLKMGHTVTVNNTAEKCDLFIQF-GARLGRTPAEVYSTCDITF 328
Db 2 KVGFIGLIGMGFPNASLLKAGYDLTVNRTIEKAEKLGKMGAKVAHSPREVAEVSIVI 61
QY 329 ACVSDPKAAKDLVLGSPGVLGIRPGKCYVDMSTVDADTVTELAQVIVSRGRFLAPVS 388
Db 62 SWVTADPDEVIFGEGVVKSNKRGILFVDMSTNSPEFAKKVTKRLSEYGEFLDAPV 121
QY 389 SGNQQLSNDGMLVILAAAGDRGLYEDSCSCFQAMGKTSFFLGEVGNAAKMWLVNVMQGSF 448
Db 122 GGDGKAREGLTITMVGKEDVFKRVEPIFKAMGNIIHVGVDVSGQALKLCNQVVVALNM 181
QY 449 ATTAEGTLIAHVTGQSQOTLLDILNQGLASIFLDQKQNLQGNFKPDFYLYKIQKDLR 508
Db 182 VSVVEGLLARSLSIGDDDKLFSVLSTGAANSFTVQYVLPKIMKGLNPGFKAHLKXDLK 241
QY 509 LALALGDAVNHPTPMAAANEVYKRAKALDQSDNDMSAVTRAY 551
Db 242 YAMEIANSKSLPLGLTSLALQLYNAMVSLGIGELGTQGLVKVY 284

RESULT 16
Q9SZE1

```


ID	Q9SZE1	PRELIMINARY;	PRT;	334 AA.	
AC	Q9SZE1				
DT	01-MAY-2000	(TRENBLrel. 13, Created)			
DT	01-MAY-2000	(TRENBLrel. 13, Last sequence update)			
DT	01-OCT-2003	(TRENBLrel. 25, Last annotation update)			
DE	Hypothetical protein (Putative dehydrogenase).				
GN	F19B15.150 OR AT4G29120.				
OS	Arabidopsis thaliana (Mouse-ear cress).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;				
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsis.				
OX	NCBI_TaxID=3702;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Bevan M., Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoeft A.,				
RA	Bancroft I., Mewes H.W., Mayer K.F.X., Lemcke K., Schueller C.;				
RL	Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RA	EU Arabidopsis sequencing project;				
RL	Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RA	Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoeft A.,				
RA	Mewes H.W., Lemcke K., Mayer K.F.X.;				
RL	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RA	EU Arabidopsis sequencing project;				
RL	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RA	Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,				
RA	Goldsmith A.D., Lee J.M., Quach H.L., Tang C., Toriumi M., Yu G.,				
RA	Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J.,				
RA	Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Koeseima E.,				
RA	Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M.,				
RA	Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A.,				
RA	Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;				
RT	"Full length cDNA of gene AT4g29120 (GI:7269810)."				
RL	Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.				
RN	[6]				
RP	SEQUENCE FROM N.A.				
RA	Yamada K., Banh J., Banno F., Chang E., Dale J.M., Goldsmith A.D.,				
RA	Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C.,				
RA	Yamanura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,				
RA	Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,				
RA	Kawai J., Kim C., Koeseima E., Lam B., Lin J., Meyers M.C., Miranda M.,				
RA	Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,				
RA	Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,				
RA	Theologis A.;				
RT	"Full length cDNA of gene AT4g29120 (GI:7269810)."				
RL	Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AL078470; CAB43926.1; -				
DR	EMBL; AL161574; CAB79670.1; -				
DR	EMBL; AF370197; AAK4012.1; -				
DR	EMBL; AY062952; AAL33784.1; -				
DR	PIR; T08967; T08967.				
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.				
DR	GO; GO:0004616; F:phosphoglucanate dehydrogenase (decarboxyla. .; IEA.				
DR	GO; GO:0008152; P:metabolism; IEA.				
DR	GO; GO:0006098; P:pentose-phosphate shunt; IEA.				
DR	InterPro; IPR006183; 6PGD.				
DR	InterPro; IPR006115; 6PGD NAD.				
DR	InterPro; IPR002227; Tyrosinase.				
DR	Pfam; PF03446; NAD binding 2; 1.				
DR	PRINTS; PR00076; 6PGDHDRGNASE.				
DR	PROSITE; PS00498; TYROSINASE_2; 1.				
KW	Hypothetical protein.				
SQ	SEQUENCE 334 AA; 35371 MW; EC0038096DB836BA CRC64;				
Query Match	13.4%; Score 384; DB 10; Length 334;				
Best Local Similarity	32.2%; Pred. No. 7.4e-19;				
Matches	85; Conservative	55; Mismatches	143; Indels	0; Gaps	0;
QY	269	KIGFLGLMGSGIVSNLLKMGHTVTVMNRTAEKCDLFIQEGARLGTPTAEVSTCDITF	328		
Db	7	KIGFIGTVMGKGNKMLKGGYTVHYVNTKEKALPLVEIGAVLETSEKLSQRCDVII	66		
QY	329	ACVSDPEKAAKDLVLGSPGVLTQGIIRPGKCYVDMSTVDADTVELAQIVISRCGRLEAPVS	388		
Db	67	TWGYPSDVEEVFNQNGILNLLKPKQSYVVDVMTTSKPSLAKKIYAAAKERNIFALDAPVS	126		
QY	389	GNQOLSDNGMLVLAAGDRGLYEDCSSCFQAMGKTSFFLGEVGNVQVQSF	448		
Query Match	13.1%; Score 375; DB 16; Length 292;				
Best Local Similarity	30.0%; Pred. No. 2.6e-18;				
Matches	85; Conservative	55; Mismatches	143; Indels	0; Gaps	0;
Matches	96; Conservative	48; Mismatches	152; Indels	2; Gaps	2;
QY	254	ADSTAVNGSIPTDKKIGFLGLMGSGIVSNLLKMGHTVTVMNRTAEKCDLFIQEGARL	313		
Db	24	ASSIISDIIPSNWKIGWIGVWGRSMCGHLIKAGYTVVFNRTTSKATQTLIDMGANV	83		
QY	314	GRTPAEVSTCDITFACVSDPKAAKDLVLGP-SGVLOGIRPGKCYVDMSTVDADTVELA	372		
Db	84	ADSPNSVAEQSDVWFTIVGYPDSVHRVLLDPKSGALSLRQGGVLVDMTTSEPSLAERIA	143		
QY	373	QVIVSRGRFLEAPVSGNQQLSNDQMLVLAAGDRGLYEDCSSCFQAMGKTSFFLGEVGN	432		
Db	144	KAASFKNCFSDAPVSGDGLGAKNKGKLSIFAGGDETTVKRLDPLFSLMGKNV-FMGTSK	202		
QY	433	AAKMLLVNMQSGSPMATIAEGLTLAHTVGTGSOQTLLDILNQQLASIFLDQKCNILQG	492		
Db	203	GQFAXLANQITASTMLGIVEGLIYAHKAGLDVKKFLEAISTGAAGSKSIDLYGRILKR	262		
QY	493	NFKPDFYLYKIQOLRLAIALGDVAVNHPTPMAAANEVYKRAKALDOSNDMSAVYRA	550		
Db	263	DFDPGFYVNHFKVLDLIGICLNQCQRMGLAPGLALAQQLYLSLKAHGEGLGTQALLA	320		
RESULT 17					
Q97DX8					
ID	Q97DX8	PRELIMINARY;	PRT;	292 AA.	
AC	Q97DX8				
DT	01-OCT-2001	(TRENBLrel. 18, Created)			
DT	01-OCT-2001	(TRENBLrel. 18, Last sequence update)			
DT	01-JUN-2003	(TRENBLrel. 24, Last annotation update)			
DE	Dehydrogenase related to 3-hydroxyisobutyrate dehydrogenase, YKWC				
DE	B.subtilis ortholog.				
GN	CAC3342.				
OS	Clostridium acetobutylicum.				
OC	Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;				
OC	Clostridium.				
OX	NCBI_TaxID=1488;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=ATCC 824 / DSM 792 / VKM B-1787;				
RX	MEDLINE=21359325; PubMed=11466286;				
RA	Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,				
RA	Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Soucaille P., Daly M.J.,				
RA	Tatunov R.L., Samathe F., Doucette-Stamm L., Smith D.R.;				
RA	Bennett G.N., Koonin E.V., Smith D.R.;				
RT	"Genome sequence and comparative analysis of the solvent-producing				
RT	bacterium Clostridium acetobutylicum."				
RT	J. Bacteriol. 183:4823-4838(2001).				
DR	EMBL; AE007831; AAK81274.1; -				
DR	PIR; G97310; G97310.				
DR	GO; GO:0004616; F:phosphoglucanate dehydrogenase (decarboxyla. .; IEA.				
DR	GO; GO:0006098; P:pentose-phosphate shunt; IEA.				
DR	InterPro; IPR006183; 6PGD.				
DR	InterPro; IPR006115; 6PGD NAD.				
DR	Pfam; PF03446; NAD binding_2; 1.				
DR	PRINTS; PR00076; 6PGDHDRGNASE.				
KW	Complete proteome.				
SQ	SEQUENCE 292 AA; 31542 MW; 7EF0ECC9D59EED87 CRC64;				

```
Db 127 GGDIGAAEGTSLIMVGGDEKFAIDALPIFELMGKNIVLQSGAGSGOHTKMCNQIAIASNI 186
QY 449 ATIAEGLTLAHVTGQSQTLLDILNOQLASIFLDQKCNILQGNKEKPDFYLVKIQKDLR 508
Db 187 MGVCALSYAKKSGINAEVLTKSGAGAAASWQLSAYAPRLKGDFFNPGFYIKHFVKDMK 246
QY 509 LAIALGDVANHPTPMAAANAENVYKRAKALDQSDNDMSAVYRAY 551
Db 247 IALBEADMGKLTFALELSKLYDKLIKEGKEDLGTQALYHLY 289

RESULT 18
Q8ZXR3 ID Q8ZXR3 PRELIMINARY; PRT; 284 AA.
AC Q8ZXR3;
DT 01-MAR-2002 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Hypothetical protein PAE1145.
GN PAE1145.
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX MEDLINE=21664397; PubMed=11792869;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Muller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989 (2002).
DR EMBL; AE009806; AAL63283.1; -.
DR GO; GO:0008442; F:3-hydroxyisobutyrate dehydrogenase activity; IEA.
DR GO; GO:0004616; F:phosphogluconate dehydrogenase (decarboxyla. . .; IEA.
DR GO; GO:0006098; P:pentose-phosphate shunt; IEA.
DR GO; GO:0006573; P:valine metabolism; IEA.
DR InterPro; IPR002204; 3hydroxisobut_dh.
DR InterPro; IPR006183; 6PGD.
DR InterPro; IPR006115; 6PGD_NAD.
DR Pfam; PF03446; NAD_binding_2; 1.
DR PRINTS; PR00076; 6EGDHRGNASE.
DR PROSITE; PS00895; 3 HYDROXYISOBUT DH; 1.
KW Hypothetical protein; Complete proteome.
SQ SSSEQUENCE 284 AA; 30155 MW; 9D845881717D8550 CRC64;

Query Match 13.0%; Score 373; DB 17; Length 284;
Best Local Similarity 30.2%; Pred. No. 3.5e-18;
Matches 80; Conservative 54; Mismatches 131; Indels 0; Gaps 0;

QY 269 KIGFLGILGMSGIVSNLLKMGHTVTVNNRTAEKCDLFIQEGARLGTAEVYSTCDITF 328
Db 2 RVGVGLGIMGGPMAMHLRAGLLAAVYNTRAKAEPEKLVGVVAGSPGLAKRVDDVI 61
QY 329 ACVSDPKAAKDLVLGSPGVLGIRPGKCYVDMSTVDADTVELAQVTSRGRFLEAPVS 388
Db 62 IMVSDAPDVEQLVFGPGGIVEGARPGLIIVDMSTNSPDWARRFAERLAQYGVFELAPVT 121
QY 389 GNOQLSNDGMVLILAGDRGLYEDCSSCFQAMGKTSFFLGEVGNAAKMLIVNNVQGSFM 448
Db 122 GGQKGAVEGTLTVMGVKEELFKRLLPVQAFGNIVYVGVPGYQAMKLVNQVVAALNT 181
QY 449 ATIAEGLTLAHVTGQSQTLLDILNOQLASIFLDQKCNILQGNKPDFYLVKIQKDLR 508
Db 182 VAMVEGRLAKALGLDMKVVQVLTGGAARSGSIELYLPKLLRGDLTPGPKAAHLKKDLS 241
QY 509 LAIALGDVANHPTPMAAANAENVYK 533
Db 242 YVMELANRASLSLPATALALELYKK 266

RESULT 19
```

```
Q89HA0 ID Q89HA0 PRELIMINARY; PRT; 313 AA.
AC Q89HA0;
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Oxidoreductase.
GN BLR6094.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=USDA 110.
RX MEDLINE=2248498; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpou S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197 (2002).
DR EMBL; AP005957; BAC51359.1; -.
DR GO; GO:0004616; F:phosphogluconate dehydrogenase (decarboxyla. . .; IEA.
DR GO; GO:0006098; P:pentose-phosphate shunt; IEA.
DR InterPro; IPR006183; 6PGD.
DR InterPro; IPR006115; 6PGD_NAD.
DR Pfam; PF03446; NAD_binding_2; 1.
DR PRINTS; PR00076; 6EGDHRGNASE.
KW Complete proteome.
SQ SEQUENCE 313 AA; 32732 MW; 3B92612722105D51 CRC64;

Query Match 13.0%; Score 372.5; DB 16; Length 313;
Best Local Similarity 31.2%; Pred. No. 4.3e-18;
Matches 93; Conservative 55; Mismatches 145; Indels 5; Gaps 3;

QY 254 ADSTAVNGSITPTDKKIGFLGILGMSGIVSNLLKMGHTVTVNNRTAEKCDLFIQEGARL 313
Db 13 ADASKAN---PQDERFGYLGILGMLGTPMTRLLKAGYQVSNRSRGKVVPVIVEAGARH 69
QY 314 GRTPAEVYSTCDITFACVSDPKAAKDLVLGSPGVLGIRPGKCYVDMSTVDADTVELA- 372
Db 70 GASPRDVMAGSDIVFMCVTDAAAEVIFGAEGLSFAPGAGKLVDFSSIFHPDAARDLAT 129
QY 373 QVIVSRGRFLEAPVSGNQOLSNDGMVLILAGDRGLYEDCSSCFQAMGKTSFFLGEVGN 432
Db 130 RLKAANGAGWIDAPVSGGTGAEGLTIAIMAGGDASDIERVVPVLTMARFRTHMGPTGA 189
QY 433 AAKMLIVNNVQGSFWATIAEGLTLAHVTGQSQTLLDILNOQLASIFLDQKCNILQ 492
Db 190 GQTAKLCNQVIGCANVLAELAEATRLAVNAGIDANRULFEALAGGFADSIPLQLFVPRVQ 249
QY 493 NFKPDF-YLKYIQKRLALALGDVANHPTPMAAANAENVYKRAKALDQSDNDMSAVYR 549
Db 250 IHSEPLGHITATMLKDLTDVADVAQTSTPVPMTATLAGQLFRLAKAARGADADALEIYK 307

RESULT 20
Q8U2W2 ID Q8U2W2 PRELIMINARY; PRT; 278 AA.
AC Q8U2W2;
DT 01-JUN-2002 (T-EMBLrel. 21, Created)
DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE 3-hydroxyisobutyrate dehydrogenase.
GN PF0716.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2261;
[1]
RN SEQUENCE FROM N.A.
```


DR	GO: 0004616; F: phosphogluconate dehydrogenase (decarboxyla. . .; IEA.
DR	GO: 0006098; P: pentose-phosphate shunt; IEA.
DR	GO: 0006573; P: valine metabolism; IEA.
DR	InterPro: IPR002204; 3hydroxisobut_dh.
DR	InterPro: IPR006183; 6PGD.
DR	InterPro: IPR006115; 6PGD_NAD.
DR	InterPro: IPR006398; tartaric sem red.
DR	Fam: PF03446; NAD_binding_2; 1.
DR	PRINTS: PR00076; 6PGDHRGNASE.
DR	TIGRFAMs: TIGR01505; tartaric sem red; 1.
DR	PROSITE: PS00895; 3_HYDROXYISOBUT_DH; 1.
KW	Complete proteome.
Q	SEQUENCE 299 AA; 31019 MW; 08E65F785BC2B566 CRC64;
Query Match	12.8%; Score 367; DB 16; Length 299;
Best Local Similarity	29.4%; Pred. No. 9,9e-18;
Matches	83; Conservative 57; Mismatches 140; Indels 2; Gaps 1;
Qy	269 KIGFLGILMGSGTVSNLLKMGHTVYVNNRTAEKCDLFIQBARLGRTPAEVWSTDTTF 328
Db	7 KVGFIGLIGMGKPSKLLKAGYSLVADRNPEAIADVAAGAETASTAKAIEQCDVII 66
Qy	329 ACVSDPKAAKDLVLGSPGVLOGIPGKCYVDMSTVDATVTTELAQVIVSRGRLEAPVS 388
Db	67 TMLPNSPHVKEVALGEGNIIIEGAKPGTFLIDMSSIAFLASREISEALKAKGIDMLDAPVS 126
Qy	389 GNOQLSNDGMLVILAAAGDRGLYEDCSSCFQAMGKTSFFLGSGVGNAAKMMLIVNMQGSM 448
Db	127 GGEPKAIDGTLVVMVGDKAIFDKYDLMKAKAGSVVHTGEIGAGNVTKLANQVIVALNI 186
Qy	449 ATIAEGHTLAHVTGOSOOTLDLILNQGLASIFLDQKQCNLTQGNFKPDFYLYIQKDLR 508
Db	187 AAMSEALTATKAGWPNPLVQAIRGLLAGSTVLDAKAPWMDNRNFKPGPRIDLHDKLA 246
Qy	509 LAIALGDAVNHPTMAAAAEVYK--RAKALDQSDNDMSAVY 548
Db	247 NALDTSHGVAQLPETAAVMEMMQALRADGLGTADHSALACY 288
RESULT 23	
Q8ZLV8	PRELIMINARY; PRT; 296 AA.
ID	Q8ZLV8
AC	Q8ZLV8; 20, Created
DT	01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT	01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE	Tartronate semialdehyde reductase (TSAR) (EC 1.1.1.60).
GN	GARR OR SIM3248.
OS	Salmonella typhimurium.
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC	Enterobacteriaceae; Salmonella.
OX	NCBI_TaxID=602;
XP	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=LT2 / SGSC1412 / ATCC 700720;
FX	MEDLINE=21534948; PubMed=11677609;
RA	McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA	Courtney L., Porwollik S., Ali J., Pante M., Du F., Hou S., Layman D.,
RA	Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA	Ryan B., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA	Waterston R., Wilson R.K.;
RT	"Complete genome sequence of Salmonella enterica serovar Typhimurium
RT	LT2.";
RL	Nature 413:852-856(2001).
DR	EMBL; AE008849; AAU22120.1; -.
DR	GO: 0008679; F:2-hydroxy-3-oxopropionate reductase activity; IEA.
DR	GO: 0008442; F:3-hydroxyisobutyrate dehydrogenase activity; IEA.
DR	GO: 0016491; F:oxidoreductase activity; IEA.
DR	GO: 0004616; F:phosphogluconate dehydrogenase (decarboxyla. . .; IEA.
DR	GO: 0006098; P: pentose-phosphate shunt; IEA.
DR	GO: 0006573; P: valine metabolism; IEA.
DR	InterPro: IPR002204; 3hydroxisobut_dh.
DR	InterPro: IPR006183; 6PGD.


```

KW Complete proteome.
SQ SEQUENCE 301 AA; 31326 MW; 91AC4FBF73F6925F CRC64;

Query Match 12.4%; Score 354.5; DB 16; Length 301;
Best Local Similarity 29.6%; Pred. No. 7.6e-17;
Matches 87; Conservative 56; Mismatches 134; Indels 15; Gaps 6;

QY 264 TPTDKKTGFLGLGMSGIVSNLLKMGHTVTVMNRTAEKCDLFIQE-GARLGRTPAEVS 322
Db 11 TTTTPTNVAFLGLGVMGLPMAGHLARAGHQVTYVNRATAQAQAARAEFGKAAATPREAAA 70

QY 323 TCDITFACVSDPRAAKDLVLGSPGFLQGIIRPKCYVDMSTVDADTVELAQIVSRGGRF 382
Db 71 GAQIVFCVCGNDDRLRSVLTGADGAFAGMQAGAVFVDHTTASADVAARELYAARELMLQF 130

QY 383 LEAPVSGNQQLSNDMLVILAAAGDRGLYEDCSSCFQAMGKTSFFLGE--VGNAAKM--M 437
Db 131 IDAPVSGQAGAVNGALTIMCGSDSATFEDIKPAVAHAFRAVTLVGPFGAGOLAKMNQV 190

QY 438 LIVNMGVSFWATTIARGTLTAHVTGQSOOTLLDILNQOLASIFLDQKCCNQLQGNFKPD 497
Db 191 CIAGVVOG-----LSEAIAPQQTAGLDMLVLDVISKGAQSQWMENRGATMVDQDFG 245

QY 498 FYLXVIOKLELATALGDAVNHPT--PMAAANAENVYKAKALDOSDNDMSAVYR 549
Db 246 FAVDMRMDGLGVIA--EARNNGARLPLTALVDQFYGDVQKMGNNRWDTSLLIK 297

RESULT 29
QESP3 PRELIMINARY; PRT; 298 AA.
AC Q8ESP3
AC Q8ESP3;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DE 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE 3-hydroxysobutyrate dehydrogenase (EC 1.1.1.31).
GN CB0581.
OS Oceanobacillus theyensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.
NCBI TaxID=182710;
RN [1]
RX SEQUENCE FROM N.A.
RC STRAIN-HT831 / DSM 14371 / JCM 11309;
RX MEDLINE=22220767; PubMed=12235376;
RA Takami H., Takaki Y., Uchiyama I.;
RT "Genome sequence of Oceanobacillus theyensis isolated from the Iheya
RT Ridge and its unexpected adaptive capabilities to extreme
RT environments.";
RL Nucleic Acids Res. 30:3927-3935(2002).
DR EMEL; AP004595; BAC12537.1; --
DR GO; GO:0008442; F:3-hydroxysobutyrate dehydrogenase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0004616; F:phosphogluconate dehydrogenase (decarboxyla. .; IEA.
DR GO; GO:0006098; P:pentose-phosphate shunt; IEA.
DR InterPro; IPR006183; 6PGD.
DR InterPro; IPR006115; 6PGD NAD.
DR Pfam; PF03446; NAD_binding_2; 1.
DR PRINTS; PR00076; 6PGDHRGNASE.
DR Oxidoreductase; Complete proteome.
KW SEQUENCE 298 AA; 31931 MW; 4D80E5EAAF7C374 CRC64;

Query Match 12.3%; Score 353; DB 16; Length 298;
Best Local Similarity 27.4%; Pred. No. 9.5e-17;
Matches 77; Conservative 56; Mismatches 148; Indels 0; Gaps 0

QY 270 IGFLGLGLMSGIVSNLLKMGHTVTVMNRTAEKCDLFIQEGARLGRTPAEVSTCDITPA 329
Db 9 IGFIGTVGVGKSMKXNLLNAGYSIRIFTRTKSKADLLNIGAKWDKDIPTLAKQCDVIT 68

QY 330 CVSDPRAAKDLVLGSPGVLOGIRPKCYVDMSTVDADTVELAQIVSRGGRFLEAPVSG 389
Db 69 MVGVPDSVEVYLGEGLINKCNKGSYLIDMTTSKPSLAEEIFSVSEKGLHVLDPVSG 128

```


Search completed: September 16, 2004, 07:32:08
Job time : 54 secs

Title: GenCore version 5.1.6
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2004, 07:30:42 ; Search time 55 Seconds
 (without alignments)
 195.215 Million cell updates/sec

Title: US-10-067-482-2_COPY_271_308
 Perfect score: 200
 Sequence: 1 GFLGLMGSGIVSNLLKMGHTVTWNRKTAERCDLFIQ 38

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : A_Geneseq_29Jan04.*

- 1: Geneseqp1980s.*
- 2: Geneseqp1990s.*
- 3: Geneseqp2000s.*
- 4: Geneseqp2001s.*
- 5: Geneseqp2002s.*
- 6: Geneseqp2003as.*
- 7: Geneseqp2003bs.*
- 8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	200	100.0	472	6	Aae38240 Human enz
2	200	100.0	553	6	AAO16415 Human nuc
3	166	83.0	547	2	Aaw69240 Clone AQ7
4	166	83.0	547	4	Aab90720 Human AQ7
5	166	83.0	550	4	AAU21690
6	166	83.0	550	7	ADC46331
7	126	63.0	574	4	ABE60930
8	125	62.5	161	7	ABM74242 Drosophil
9	125	62.5	168	3	AAU41222
10	125	62.5	289	3	AAH19491
11	125	62.5	289	6	AAE36392
12	124	62.0	345	5	AAE19929
13	123	61.5	343	3	AAE10424
14	123	61.5	353	3	AAE10423 Arabidops
15	118	59.0	290	6	AAE36391
16	117	58.5	360	5	AAE19927
17	109	54.5	292	6	ABU2159
18	106	53.0	288	6	ABU17331
19	106	53.0	299	6	ADA36120
20	105	52.5	290	6	ABU17066
21	102	51.0	326	6	ADA36852
22	101	50.5	285	3	AAE10425
23	100	50.0	295	6	ABU29967
24	100	50.0	299	7	ADC95482 E. faeciu
25	93	46.5	492	2	AAW27613 Brevibact

26	92	46.0	158	3	AAO6417	Aag06417 Arabidops
27	92	46.0	180	3	AAO6416	Aag06416 Arabidops
28	92	46.0	182	3	AAO6415	Aag06415 Arabidops
29	92	46.0	312	3	AAO20956	Aag20956 Arabidops
30	92	46.0	334	3	AAO20955	Aag20955 Arabidops
31	92	46.0	336	3	AAO20954	Aag20954 Arabidops
32	91.5	45.8	67	4	ABE67146	ABE67146 Drosophil
33	91	45.5	158	3	AAO41223	Aag41223 Arabidops
34	89	44.5	296	6	ABU28970	ABU28970 Protein e
35	88	44.0	464	6	ABU19266	ABU19266 Staphyloc
36	88	44.0	465	4	AAU34285	Aau34285 Staphyloc
37	88	44.0	468	4	AAU36997	Aau36997 Staphyloc
38	88	44.0	468	6	ABM71061	ABM71061 Staphyloc
39	87	43.5	324	4	ABE70141	ABE70141 Drosophil
40	87	43.5	468	6	ABU15983	ABU15983 Protein e
41	86	43.0	459	4	AAO79095	Aag79095 Amino aci
42	86	43.0	472	5	ABE48294	ABE48294 Listeria
43	86	43.0	472	6	ABU32435	ABU32435 Protein e
44	86	43.0	479	4	AAE79369	AAE79369 Coryneb
45	86	43.0	484	4	AAE79368	AAE79368 Coryneb
46	86	43.0	492	4	AAO91351	Aag91351 C glutam
47	86	43.0	492	7	ADC35705	Adc35705 6-phospho
48	86	43.0	492	7	ADD13347	Add13347 C. glutam
49	85	42.5	204	2	AAV34958	Aay34958 Chlamydia
50	85	42.5	479	6	ABU26880	ABU26880 Protein e
51	85	42.5	479	6	ABU27221	ABU27221 Protein e
52	85	42.5	484	6	ABU26044	ABU26044 Protein e
53	84.5	42.2	299	3	AAE15899	AAE15899 E. coli p
54	84	42.0	299	5	ABE91546	ABE91546 Herbicida
55	84	42.0	488	6	ABU48551	ABU48551 Protein e
56	82	41.0	154	3	AAO41224	Aag41224 Arabidops
57	82	41.0	293	5	ABE5570	ABE5570 Lactococc
58	82	41.0	297	5	ABE91547	ABE91547 Herbicida
59	82	41.0	469	6	ABU21852	ABU21852 Protein e
60	81	40.5	247	5	AAE19926	AAE19926 Corn 3-hy
61	81	40.5	286	5	ABE49230	ABE49230 Listeria
62	81	40.5	286	6	ABU32839	ABU32839 Protein e
63	80	40.0	254	4	ABG17598	ABG17598 Novel hum
64	80	40.0	439	6	ABU34614	ABU34614 Protein e
65	80	40.0	484	4	AAU35461	Aau35461 Haemophil
66	80	40.0	484	6	ABU30295	ABU30295 Protein e
67	80	40.0	485	4	AAE81157	AAE81157 Mycobacte
68	80	40.0	485	6	ABU36691	ABU36691 Protein e
69	80	40.0	567	4	AAU39521	AAU39521 Propionib
70	80	40.0	567	6	ABM36040	ABM36040 Propionib
71	79	39.5	292	4	ABE52479	ABE52479 Escherich
72	79	39.5	322	2	AAV37157	AAV37157 Protein i
73	79	39.5	474	6	ABU43488	ABU43488 Protein e
74	79	39.5	484	6	ABU39412	ABU39412 Protein e
75	77	38.5	92	3	AAO12290	Aag12290 Zea mays
76	77	38.5	96	3	AAO26084	Aag26084 Zea mays
77	77	38.5	113	3	AAO19126	Aag19126 Zea mays
78	77	38.5	470	6	ABU19468	ABU19468 Protein e
79	77	38.5	471	6	ABU22158	ABU22158 Protein e
80	77	38.5	479	7	ABM73851	ABM73851 DNA clone
81	77	38.5	480	6	ADA48242	Ada48242 Rice prot
82	76	38.0	468	3	AAV99866	AAV99866 Escherich
83	76	38.0	482	6	ABU49748	ABU49748 Protein e
84	76	38.0	483	6	ABU36070	ABU36070 Protein e
85	76	38.0	489	6	ABU33961	ABU33961 Protein e
86	75	37.5	347	3	AAO52190	Aag52190 Arabidops
87	75	37.5	347	3	AAO8461	Aag08461 Arabidops
88	75	37.5	468	4	AAO82286	Aag82286 S. epider
89	75	37.5	468	6	ABU42697	ABU42697 Protein e
90	75	37.5	474	5	ABE38396	ABE38396 Staphyloc
91	74	37.0	302	6	ABU21583	ABU21583 Protein e
92	74	37.0	461	6	ABU31389	ABU31389 Protein e
93	74	37.0	467	6	ABU45660	ABU45660 Protein e
94	74	37.0	468	3	AAV99857	AAV99857 Escherich
95	74	37.0	468	4	AAU38229	Aau38229 Salmonell
96	74	37.0	468	6	ABU48164	ABU48164 Protein e
97	74	37.0	486	3	AAO16834	Aag16834 Arabidops
98	74	37.0	521	3	AAO16833	Aag16833 Arabidops

99 73 36.5 192 3 AAY99862 Aay99862 Escherich
100 73 36.5 298 7 ADD95220 Add95220 Alcaligen

ALIGNMENTS

RESULT 1
AAE38240
ID AAE38240 standard; protein; 472 AA.
AC AAE38240;
XX
XX
XX 20-NOV-2003 (first entry)
XX
XX Human enzyme (ENZM) protein #32.
XX Human; enzyme; ENZM; immune disorder; infection; myocardial infarction;
KW gene therapy; anaemia; acquired immune deficiency syndrome; infection;
KW reproductive disorder; cardiovascular; eye; cell proliferation; cancer;
KW AIDS; allergy; asthma; Addison's disease; diabetes; goitre; impotence;
KW infertility; atherosclerosis; metabolic disorder.
XX
XX Homo sapiens.
OS
XX
XX WO2003052075-A2.
PN
XX
XX 26-JUN-2003.
PD
XX
XX 12-DEC-2002; 2002WO-US040161.
PF
XX
XX 14-DEC-2001; 2001US-0340357P.
PR
XX 20-DEC-2001; 2001US-0342962P.
PR
XX 21-DEC-2001; 2001US-0343558P.
PR
XX 22-JAN-2002; 2002US-0351107P.
PR
XX
XX (INCY-) INCYTE GENOMICS INC.
PA
XX Chawla NK, Lee SY, Ring HZ, Lee EA, Forsythe IJ, Khare R;
PI Tran UK, Kable AE, Richardson TW, Emerling BM, Lindquist EA;
PI Baughn MR, Hafalia AJA, Jin P, Swarnakar A, Li JX, Marquis JP;
PI Lee S, Gorvad AE, Sprague WW, Becha SD, Elliott VS;
XX
XX WPI; 2003-533016/50.
DR
XX N-PSDB; AAD57520.
DR
XX
XX New human enzyme (ENZM) polypeptide, useful for preparing a composition
PT for treating a disease associated with decreased expression or
PT overexpression of ENZM e.g. cancer.
XX
XX Claim 1; Page 298-299; 264pp; English.
XX
XX The invention relates to human enzyme (ENZM) polypeptides and their
CC corresponding polynucleotides. ENZM sequences are useful for preparing a
CC composition for diagnosing or treating a disease or condition associated
CC with decreased expression or overexpression of functional ENZM. The
CC disorders include immune disorders (anaemia, allergy or asthma),
CC infectious disorders (viral, fungal, parasitic or protozoal infection),
CC immune deficiencies (acquired immune deficiency syndrome; AIDS),
CC metabolic disorder (Addison's disease, diabetes or goitre), reproductive
CC disorders (infertility or impotence), cardiovascular disorders
CC (atherosclerosis or myocardial infarction), eye disorders and cell
CC proliferative disorders (cancer). ENZM DNA is useful in gene therapy. The
CC present sequence is human ENZM protein
XX
XX Sequence 472 AA;
SQ

Query Match 100.0%; Score 200; DB 6; Length 472;
Best Local Similarity 100.0%; Pred. No. 4.3e-20;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GFLGLGLMGSGIVSNLLKMGHTVTWNRATKCDLFIQ 38
|||||

Db 190 GFLGLGLMGSGIVSNLLKMGHTVTWNRATKCDLFIQ 227

RESULT 2

AAO16415
ID AAO16415 standard; protein; 553 AA.
XX
XX
XX AAO16415;
AC
XX
XX 10-APR-2003 (first entry)
DT
XX
XX Human nucleic acid-associated protein (NAAP) - SEQ ID No 12.

XX
XX Human; nucleic acid-associated protein; NAAP; arteriosclerosis;
KW cell proliferative disorder; atherosclerosis; cirrhosis; hepatitis; AIDS;
KW cancer; developmental disorder; renal tubular acidosis; anaemia; asthma;
KW mental retardation; neurological disorder; Alzheimer's disease; epilepsy;
KW Parkinson's disease; autoimmune disorder; inflammatory disorder; allergy;
KW Crohn's disease; transgenic animal; animal model.
XX
XX Homo sapiens.
OS
XX
XX WO2003000864-A2.
PN
XX
XX 03-JAN-2003.
PD
XX
XX 20-JUN-2002; 2002WO-US021179.
PF
XX
XX 22-JUN-2001; 2001US-0300518P.
PR
XX 29-JUN-2001; 2001US-0301787P.
PR
XX 29-JUN-2001; 2001US-0301792P.
PR
XX 29-JUN-2001; 2001US-0301892P.
PR
XX 06-JUL-2001; 2001US-0301893P.
PR
XX 06-JUL-2001; 2001US-0303405P.
PR
XX 06-JUL-2001; 2001US-0303442P.
PR
XX 15-MAR-2002; 2002US-0364438P.
PR
XX
XX (INCY-) INCYTE GENOMICS INC.

XX
XX Gandhi AR, Swarnakar A, Hafalia AJA, Warren BA, Emerling BM;
PI Arvizu CS, Ison CH, Honchell CD, Lee EA, Yue H, Forsythe IJ;
PI Ramkumar J, Griffin JA, Yang J, Sanjanwala MM, Baughn MR;
PI Borowsky ML, Yao MG, Walia NK, Bandnan O, Lal PG, Becha SD, Lee SY;
PI Richardson TW, Elliott VS, Luo W, Tang Y, Zebbarjadian Y, Lu Y;
XX
XX WPI; 2003-201420/19.
DR
XX N-PSDB; AAL51565.
DR
XX
XX New nucleic acid-associated proteins and polynucleotides, useful for
PT diagnosing, treating or preventing cell proliferative (e.g. cancer),
PT neurological (e.g. epilepsy or Parkinson's disease), or autoimmune
PT disorders (e.g. AIDS).
PT
PT
PT
XX
XX Claim 1; Page 227-228; 312pp; English.

XX
XX The invention comprises the amino acid and coding sequences of human
CC nucleic acid-associated proteins (NAAP). The DNA and protein sequences of
CC the invention are useful for diagnosing, treating or preventing disorders
CC associated with aberrant expression of NAAP, such as: cell proliferative
CC disorders (e.g. arteriosclerosis, atherosclerosis, cirrhosis, hepatitis
CC or cancer); developmental disorders (e.g. renal tubular acidosis, anaemia
CC or mental retardation); neurological disorders (e.g. Alzheimer's disease,
CC Parkinson's disease or epilepsy); and autoimmune/inflammatory disorders
CC (e.g. AIDS, allergies, asthma or Crohn's disease). The DNA sequences of
CC the invention are useful for creating transgenic animals to model human
CC disease. The present amino acid sequence represents a human nucleic acid-
CC associated protein of the invention
XX
XX Sequence 553 AA;
SQ

Query Match 100.0%; Score 200; DB 6; Length 553;
Best Local Similarity 100.0%; Pred. No. 5.2e-20;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GFLGLMGSGIVSNLLKMGHTVTWNRRTAEKCDLFQ 38
 DB 271 GFLGLMGSGIVSNLLKMGHTVTWNRRTAEKCDLFQ 308

RESULT 3
 AAW69240
 ID AAW69240 standard; protein; 547 AA.
 AC AAW69240;
 XX
 DT 21-OCT-1998 (first entry)
 XX
 DE Clone AQ73_3 protein sequence.
 XX
 KW Secreted protein; nutritional source; cell proliferation activity;
 KW cell differentiation activity; immune stimulant; tissue growth activator;
 KW haematopoiesis regulator; anti-inflammatory; tumour invasion suppressor;
 KW tumour inhibitor; clone AQ73_3.
 XX
 OS Homo sapiens.
 XX
 PN WO9825962-A2.
 XX
 PD 18-JUN-1998.
 XX
 PF 12-DEC-1997; 97WO-US023224.
 XX
 PR 13-DEC-1996; 96US-00766263.
 PR 11-DEC-1997; 97US-00989232.
 XX
 PA (GEMY) GENETICS INST INC.
 XX
 PI Jacobs K, McCoy JM, Lavallie ER, Racie LA, Merberg D, Treacy M;
 PI Spaulding V, Agostino MJ;
 XX
 DR WPI; 1998-362424/31.
 DR N-PSDB; AAV44859.
 XX
 PT New isolated polynucleotides - obtained from human adult testis, human
 PT adult ovary, human adult brain and human adult heart cDNA libraries.
 XX
 PS Claim 15; Page 66-67; 108pp; English.
 XX
 CC This sequence is secreted protein, encoded by a polynucleotide of the
 CC invention. The DNA was isolated from a human adult ovary cDNA library,
 CC and is designated clone AQ73_3. The DNA sequences and encoded
 CC polypeptides can be used as nutritional sources or supplements, or may
 CC exhibit e.g. cytokine and cell proliferation/differentiation activity,
 CC immune stimulating or suppressing activity, haematopoiesis regulating
 CC activity, receptor/ligand activity, anti-inflammatory activity,
 CC activin/inhibin activity, chemostatic/chemokinetic activity,
 CC cadherin/tumour invasion suppressor activity, tissue growth activity,
 CC tumour inhibition activity or other activities
 XX
 SQ Sequence 547 AA;
 Query Match 83.0%; Score 166; DB 2; Length 547;
 Best Local Similarity 100.0%; Pred. No. 4.3e-15;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GFLGLMGSGIVSNLLKMGHTVTWNRRTAEK 32
 DB 271 GFLGLMGSGIVSNLLKMGHTVTWNRRTAEK 302

RESULT 4
 AAB90720
 ID AAB90720 standard; protein; 547 AA.
 AC AAB90720;
 XX
 XX

DT 07-JUN-2001 (first entry)
 DE Human AQ73_3 protein sequence SEQ ID 137.
 XX
 KW Human; secreted protein; nutrient; cytokine modulator; proliferation;
 KW differentiation; immune system modulator; tissue growth; chemotactic;
 KW haemostatic; thrombolytic; anti-inflammatory; tumour inhibition;
 KW haematopoiesis.
 XX
 OS Homo sapiens.
 XX
 PN WO200119988-A1.
 XX
 PD 22-MAR-2001.
 XX
 PF 14-SEP-2000; 2000WO-US025135.
 XX
 PR 17-SEP-1999; 99US-00398829.
 XX
 PA (GEMY) GENETICS INST INC.
 XX
 PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
 PI Merberg D, Treacy M, Bowman MR, Spaulding V, Agostino MJ;
 XX
 DR WPI; 2001-244801/25.
 DR N-PSDB; AAF98456.
 XX
 PT Isolated nucleic acids encoding polypeptides, useful for modulating e.g.
 PT cytokine and cell proliferation/differentiation activity, the immune
 PT system and hematopoiesis regulating activity.
 XX
 PS Disclosure; Page 468-470; 557pp; English.
 XX
 CC Human cDNA clones represented in AAF98374 - AAF98489 encode secreted
 CC proteins AAB90667 - AAB90750. The cDNA clones are isolated from various
 CC tissue types, and may be used in the prevention, treatment and diagnosis
 CC of diseases associated with inappropriate protein expression. The
 CC polypeptides and nucleic acids may be used as nutrients or to modulate
 CC cytokine and cell proliferation/differentiation activity and may also be
 CC involved in modulation of the immune system. The cDNA sequences,
 CC proteins, their agonists and/or antagonists exhibit haematopoiesis
 CC chemotactic/chemokinetic activity; haemostatic and thrombolytic activity;
 CC receptor/ligand activity; anti-inflammatory activity; activin/inhibin activity;
 CC activity; cadherin/tumour suppressor activity; and/or tumour inhibition
 CC activity. Included in the invention are probes represented in AAF98490 -
 CC AAF98572 which are specific for the cDNA clones encoding the secreted
 CC proteins
 XX
 SQ Sequence 547 AA;
 Query Match 83.0%; Score 166; DB 4; Length 547;
 Best Local Similarity 100.0%; Pred. No. 4.3e-15;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GFLGLMGSGIVSNLLKMGHTVTWNRRTAEK 32
 DB 271 GFLGLMGSGIVSNLLKMGHTVTWNRRTAEK 302

RESULT 5
 AAU21690
 ID AAU21690 standard; protein; 550 AA.
 XX
 AC AAU21690;
 XX
 DT 04-DEC-2001 (first entry)
 XX
 DE Novel human neoplastic disease associated polypeptide #123.
 XX
 KW Human; neoplastic disease associated polypeptide; cancer;
 KW hyperproliferative disorder; neural disorder; immune system disorder;
 KW muscular disorder; reproductive disorder; gastrointestinal disorder;

PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
PA Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-465558/50.
XX N-PSDB; AAS34889.
XX Novel polypeptides and polynucleotides useful as diagnostic reagents to
PT diagnose diseases or disorders associated with aberrant expression or
PT activity of polypeptides, and for treating cancers, rheumatoid arthritis.
XX Claim 11; SEQ ID NO 417; 687pp; English.
XX The present invention relates to the isolation of novel human neoplastic
CC disease associated polypeptides, and cDNA (AAS34767-AAS35050) and DNA
CC sequences encoding for these polypeptides. The sequences of the invention
CC are useful in the diagnosis, treatment, prevention and/or prognosis of
CC disorders involving neoplastic disease such as hyperproliferative
CC disorders (e.g. leukaemia, bone cancer, bladder cancer, brain stem
CC glioma, adult liver cancer, childhood cerebellar astrocytoma, or
CC Hodgkin's lymphoma). The sequences of the invention may also be useful
CC for treating other disorders such as neural disorders, immune system
CC disorders, muscular disorders, reproductive disorders, gastrointestinal
CC disorders, pulmonary disorders, cardiovascular disorders and renal
CC disorders. The polynucleotide sequences of the invention are also useful
CC in gene therapy. AAU21568-AAU21851 represent the novel human neoplastic
CC disease associated polypeptides of the invention. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX Sequence 550 AA;
SQ

Query Match 83.0%; Score 166; DB 4; Length 550;
Best Local Similarity 100.0%; Pred. No. 4.3e-15;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GFLGLGMLGSGIVSNLLKMGHTVTVMNRTAEK 32
|||||
DB 274 GFLGLGMLGSGIVSNLLKMGHTVTVMNRTAEK 305
|||||

RESULT 6
ADC46331
ID ADC46331 standard; protein; 550 AA.
AC ADC46331;
XX
XX 18-DEC-2003 (first entry)
DT
XX Human neoplastic disease-associated gene 123 protein #1.
DE
XX Neoplastic disease-associated polypeptide; gene therapy;
XX hyperproliferative disease; cancer; autoimmune disorder; diabetes;
KW rheumatoid arthritis; systemic lupus erythematosus; multiple sclerosis;
KW autoimmune thyroiditis; haemolytic anaemia; haematopoietic disorder;
KW haematologic disorder; anaemia; thrombocytopaenia; allergic reaction;
KW asthma; eczema; inflammatory disorder; ischaemia-reperfusion injury;
KW inflammatory bowel disease; Crohn's disease; neurodegenerative disorder;
KW Alzheimer's disease; Parkinson's disease; renal disease;
KW acute glomerulonephritis; end-stage renal disease;
KW cardiovascular disorder; atherosclerosis; myocarditis;
KW infectious disease; AIDS; cachexia; anorexia; wound healing;
KW epithelial cell proliferation; Human.
XX Homo sapiens.
XX
XX US2003082758-A1.
XX
XX 01-MAY-2003.
XX

PF 22-MAR-2002; 2002US-00103313.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.

PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 13-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0246417P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764854.
(HUMA-) HUMAN GENOME SCI INC.
PR Rosen CA, Ruben SM, Barash SC;
PI WPI; 2003-786918/74.
XX N-PSDB; ADC46047.
DR
XX

XX New isolated human neoplastic disease-associated polypeptides and
PT polynucleotides, useful for diagnosing, preventing, prognosticating or
PT treating medical conditions such as cancer, AIDS, diabetes or Parkinson's
PT disease.
XX
PS Claim 11; SEQ ID NO 417; 302pp; English.
XX
XX The invention relates to one of 238 disclosed human neoplastic disease-
CC associated polypeptides encoded by 171 disclosed cDNA sequences
CC (including their domains, epitopes, full-length proteins, allelic variants
CC or species homologues). Also included are there encoding nucleic acids, a
CC recombinant vector comprising the nucleic acid, a recombinant host cell
CC comprising the nucleic acid (expressing the protein), an isolated
CC antibody that binds specifically to the isolated polypeptide, preventing,
CC treating or ameliorating a medical condition, diagnosing a pathological
CC condition or a susceptibility to a pathological condition in a subject,
CC identifying a binding partner to the polypeptide, identifying an activity
CC in a biological assay, and the gene corresponding to the cDNA sequence.
CC The polypeptides, polynucleotides and antibodies are useful for
CC detecting, preventing, diagnosing, prognosticating, treating or
CC ameliorating medical conditions such as hyperproliferative diseases or
CC cancer, autoimmune disorders (e.g. diabetes, rheumatoid arthritis,
CC systemic lupus erythematosus, multiple sclerosis, autoimmune thyroiditis
CC or haemolytic anaemia), haematopoietic or haematologic disorders (e.g.
CC anaemia or thrombocytopaenia), allergic reactions including asthma or
CC eczema, inflammatory disorders (e.g. ischaemia-reperfusion injury,
Query Match 83.0%; Score 166; DB 7; Length 550;
Best Local Similarity 100.0%; Pred. No. 4.3e-15; Mismatches 0; Gaps 0;
Matches 32; Conservative 0; Indels 0;
QY 1 GFLGLMGSGIVSNLLKMGHTVTVWNRTEAK 32
Db 274 GFLGLMGSGIVSNLLKMGHTVTVWNRTEAK 305
RESULT 7
ABB60930
ID ABB60930 standard; protein; 574 AA.
XX
AC ABB60930;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 9582.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
DR N-PSDB; ABL05033.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX Disclosure; SEQ ID NO 9582; 21pp + Sequence Listing; English.
PS

Single nucleotide polymorphism sites in barley varieties and DNA sequences containing them for analysis and identification of barley varieties and production of barley transformants with desired characteristics.

Disclosure; SEQ ID XX; 284pp; Japanese.

The present invention relates to oligonucleotide clones originating in barley (*Hordeum vulgare*) which contain single nucleotide polymorphisms (SNP). The oligonucleotides may be used for analysis of SNPs among barley varieties, identification of particular varieties and genotype-phenotype analysis, isolation of specific genes and creation of new varieties by transformation of barley varieties with them and production of new barley varieties with desired properties. The present sequence represents an oligonucleotide clone sequence featured in the specification. The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published-pct-sequences

```

XX  SQ      Sequence 161 AA;
      Query Match      62.5%; Score 125; DB 7; Length 161;
      Best Local Similarity 64.7%; Fred.No. 8.1e-10;
      Matches 22; Conservative 5; Mismatches 7; Indels 0; Gaps 0
      OY  1  GFLGLGMSGIVSNLLKMGHTVTWNRRTAEKCD 34
      Db    |||||:|:|:|:|:|:|:|:|:|:|:|:|
      48  GFLGLGIMGAPWASNLKAGCDITVWNRTKSKCD 81

RESULT 9
AAG41222
ID  AAG41222 standard; protein; 168 AA.
XX  AC      AAG41222;
XX  DT      18-OCT-2000 (first entry)
XX  DE      Arabidopsis thaliana protein fragment SEQ ID NO: 51261.
XX  KW      Protein identification; signal transduction pathway; metabolic pathway;
KW      hybridisation assay; genetic mapping; gene expression control; promoter;
KW      termination sequence.
XX  OS      Arabidopsis thaliana.
XX  PN      EPI033405-A2.
XX  PP      06-SEP-2000.
XX  XX      25-FEB-2000; 2000EP-00301439.
PR  25-FEB-1999; 99US-0121825P.
PR  05-MAR-1999; 99US-0123180P.
PR  09-MAR-1999; 99US-0123548P.
PR  23-MAR-1999; 99US-0125788P.
PR  25-MAR-1999; 99US-0126264P.
PR  29-MAR-1999; 99US-0126785P.
PR  01-APR-1999; 99US-0127462P.
PR  06-APR-1999; 99US-0128234P.
PR  08-APR-1999; 99US-0128714P.
PR  16-APR-1999; 99US-0129845P.
PR  19-APR-1999; 99US-0130077P.
PR  21-APR-1999; 99US-0130449P.
PR  23-APR-1999; 99US-0130510P.
PR  28-APR-1999; 99US-0130891P.
PR  30-APR-1999; 99US-0131449P.
PR  30-APR-1999; 99US-0132048P.
PR  04-MAY-1999; 99US-0132407P.
PR  05-MAY-1999; 99US-0132484P.
PR  06-MAY-1999; 99US-0132485P.
PR  06-MAY-1999; 99US-0132486P.
PR  06-MAY-1999; 99US-0132487P.
PR  07-MAY-1999; 99US-0132863P.
PR  11-MAY-1999; 99US-0134256P.
PR  14-MAY-1999; 99US-0134218P.
PR  14-MAY-1999; 99US-0134219P.
PR  14-MAY-1999; 99US-0134221P.
PR  14-MAY-1999; 99US-0134370P.
PR  18-MAY-1999; 99US-0134768P.
PR  19-MAY-1999; 99US-0134941P.
PR  20-MAY-1999; 99US-0135124P.
PR  21-MAY-1999; 99US-0135353P.
PR  24-MAY-1999; 99US-0135629P.
PR  25-MAY-1999; 99US-0136021P.
PR  27-MAY-1999; 99US-0136392P.
PR  28-MAY-1999; 99US-0136782P.
PR  01-JUN-1999; 99US-0137222P.
PR  03-JUN-1999; 99US-0137528P.
PR  04-JUN-1999; 99US-0137502P.
PR  07-JUN-1999; 99US-0137724P.

```

us-10-067-482-2_copy_271_308.rag

Thu Sep 16 09:36:50 2004

PR	08-JUN-1999;	99US-0138094P.	PR	09-AUG-1999;	99US-0147493P.
PR	10-JUN-1999;	99US-0138540P.	PR	09-AUG-1999;	99US-0147935P.
PR	10-JUN-1999;	99US-0138847P.	PR	10-AUG-1999;	99US-0148171P.
PR	14-JUN-1999;	99US-0139119P.	PR	11-AUG-1999;	99US-0148319P.
PR	16-JUN-1999;	99US-0139452P.	PR	12-AUG-1999;	99US-0148341P.
PR	16-JUN-1999;	99US-0139453P.	PR	13-AUG-1999;	99US-0148565P.
PR	17-JUN-1999;	99US-0139452P.	PR	13-AUG-1999;	99US-0148684P.
PR	18-JUN-1999;	99US-0139454P.	PR	16-AUG-1999;	99US-0149368P.
PR	18-JUN-1999;	99US-0139455P.	PR	17-AUG-1999;	99US-0149175P.
PR	18-JUN-1999;	99US-0139456P.	PR	18-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	99US-0139457P.	PR	20-AUG-1999;	99US-0149722P.
PR	18-JUN-1999;	99US-0139458P.	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	99US-0139459P.	PR	20-AUG-1999;	99US-0149923P.
PR	18-JUN-1999;	99US-0139460P.	PR	23-AUG-1999;	99US-0149902P.
PR	18-JUN-1999;	99US-0139461P.	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999;	99US-0139462P.	PR	25-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	99US-0139463P.	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	99US-0139730P.	PR	27-AUG-1999;	99US-0151065P.
PR	18-JUN-1999;	99US-0139763P.	PR	27-AUG-1999;	99US-0151066P.
PR	18-JUN-1999;	99US-0139817P.	PR	27-AUG-1999;	99US-0151080P.
PR	22-JUN-1999;	99US-0139899P.	PR	30-AUG-1999;	99US-0151303P.
PR	23-JUN-1999;	99US-0140335P.	PR	31-AUG-1999;	99US-0151438P.
PR	23-JUN-1999;	99US-0140354P.	PR	01-SEP-1999;	99US-0151930P.
PR	23-JUN-1999;	99US-0140695P.	PR	07-SEP-1999;	99US-0152363P.
PR	24-JUN-1999;	99US-0140823P.	PR	10-SEP-1999;	99US-0153070P.
PR	28-JUN-1999;	99US-0140991P.	PR	13-SEP-1999;	99US-0153758P.
PR	29-JUN-1999;	99US-0141287P.	PR	15-SEP-1999;	99US-0154018P.
PR	30-JUN-1999;	99US-0141842P.	PR	16-SEP-1999;	99US-0154039P.
PR	01-JUL-1999;	99US-0142154P.	PR	20-SEP-1999;	99US-0154779P.
PR	01-JUL-1999;	99US-0142055P.	PR	22-SEP-1999;	99US-0155139P.
PR	02-JUL-1999;	99US-0142390P.	PR	23-SEP-1999;	99US-0155486P.
PR	06-JUL-1999;	99US-0142803P.	PR	24-SEP-1999;	99US-0155659P.
PR	08-JUL-1999;	99US-0142920P.	PR	28-SEP-1999;	99US-0156458P.
PR	09-JUL-1999;	99US-0142977P.	PR	29-SEP-1999;	99US-0156596P.
PR	12-JUL-1999;	99US-0143542P.	PR	04-OCT-1999;	99US-0157117P.
PR	13-JUL-1999;	99US-0143624P.	PR	05-OCT-1999;	99US-0157753P.
PR	14-JUL-1999;	99US-0144005P.	PR	06-OCT-1999;	99US-0157865P.
PR	15-JUL-1999;	99US-0144085P.	PR	07-OCT-1999;	99US-0158023P.
PR	16-JUL-1999;	99US-0144086P.	PR	08-OCT-1999;	99US-0158232P.
PR	16-JUL-1999;	99US-0144325P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159329P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159637P.
PR	20-JUL-1999;	99US-0144884P.	PR	14-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-0144814P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160767P.
PR	22-JUL-1999;	99US-0145085P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160815P.
PR	22-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161406P.
PR	27-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161920P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161922P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161993P.
PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.
PR	04-AUG-1999;	99US-0147302P.			
PR	05-AUG-1999;	99US-0147192P.			
PR	05-AUG-1999;	99US-0147260P.			
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.			

Query Match 62.5%; Score 125; DB 3; Length 168;
Best Local Similarity 60.5%; Pred. No. 8.5e-10;
Matches 23; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

Thu Sep 16 09:36:50 2004

us-10-067-482-2_copy_271_308.rag

```
XX 19-FEB-2002.
PD 29-JUL-1999; 99US-00364230.
PF 31-JUL-1998; 98US-0094990P.
PR (DUPO ) DU PONT DE NEMOURS & CO E I.
PR Cahoon RE, Hitz WD, Kinney AJ, Rafalski JA;
PI WPI; 2002-204621/26.
DR N-PSDB; AAD31750.
XX Isolated polynucleotide encoding 3-hydroxyisobutyrate dehydrogenase, used
PT to produce transgenic plants having an altered expression of the enzyme.
XX Claim 1; Col 33-36; 40pp; English.
PS
XX The invention relates to nucleic acid fragments encoding branched chain
CC amino acid degradation enzymes. Particularly the invention relates to 3-
CC hydroxyisobutyrate dehydrogenase, 3-hydroxymethylglutaryl CoA (HMG-CoA)
CC lyase and isovaleryl-CoA dehydrogenase polypeptides and polynucleotides
CC derived from corn, rice and wheat. Sequences of the invention are used to
CC produce transgenic plants having an altered expression of the enzyme.
CC Polynucleotides of the invention can be used as probes for physical
CC mapping of genomes. The present sequence is soybean 3-hydroxyisobutyrate
CC dehydrogenase (EC 1.1.1.31) from ses2w.pk0002.e5 clone
XX Sequence 345 AA;
SQ
Query Match 62.0%; Score 124; DB 5; Length 345;
Best Local Similarity 64.9%; Pred. No. 2.9e-09;
Matches 24; Conservative 3; Mismatches 10; Indels 0; Gaps 0;
QY 1 GFLGLGMSGIGVSNLLKMGHVTWNRRTAEKCDLFI 37
DB 55 GFLGLGMSGFMAHNLKAGVDLTVNRTKSKCDPLI 91
RESULT 13
AAG10424
ID AAG10424 standard; protein; 343 AA.
AC AAG10424;
XX
XX 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 8740.
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 23-APR-1999; 99US-0130891P.
XX 28-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 30-APR-1999; 99US-0132407P.
XX 05-MAY-1999; 99US-0132484P.
XX 06-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 07-MAY-1999; 99US-0132487P.
XX 11-MAY-1999; 99US-0132863P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 14-MAY-1999; 99US-0134370P.
XX 18-MAY-1999; 99US-0134768P.
XX 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137222P.
XX 03-JUN-1999; 99US-0137528P.
XX 04-JUN-1999; 99US-0137502P.
XX 07-JUN-1999; 99US-0137724P.
XX 08-JUN-1999; 99US-0138094P.
XX 10-JUN-1999; 99US-0138540P.
XX 10-JUN-1999; 99US-0138847P.
XX 14-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139452P.
XX 16-JUN-1999; 99US-0139453P.
XX 17-JUN-1999; 99US-0139492P.
XX 18-JUN-1999; 99US-0139454P.
XX 18-JUN-1999; 99US-0139455P.
XX 18-JUN-1999; 99US-0139456P.
XX 18-JUN-1999; 99US-0139457P.
XX 18-JUN-1999; 99US-0139458P.
XX 18-JUN-1999; 99US-0139459P.
XX 18-JUN-1999; 99US-0139460P.
XX 18-JUN-1999; 99US-0139461P.
XX 18-JUN-1999; 99US-0139462P.
XX 18-JUN-1999; 99US-0139463P.
XX 18-JUN-1999; 99US-0139750P.
XX 18-JUN-1999; 99US-0139763P.
XX 21-JUN-1999; 99US-0139817P.
XX 22-JUN-1999; 99US-0139899P.
XX 23-JUN-1999; 99US-0140353P.
XX 23-JUN-1999; 99US-0140354P.
XX 24-JUN-1999; 99US-0140695P.
XX 28-JUN-1999; 99US-0140823P.
XX 29-JUN-1999; 99US-0140991P.
XX 30-JUN-1999; 99US-0141287P.
XX 01-JUL-1999; 99US-0141842P.
XX 01-JUL-1999; 99US-0142154P.
XX 02-JUL-1999; 99US-0142055P.
XX 06-JUL-1999; 99US-0142390P.
XX 08-JUL-1999; 99US-0142803P.
XX 09-JUL-1999; 99US-0142920P.
XX 12-JUL-1999; 99US-0142977P.
XX 13-JUL-1999; 99US-0143542P.
XX 14-JUL-1999; 99US-0143624P.
XX 15-JUL-1999; 99US-0144005P.
XX 16-JUL-1999; 99US-0144085P.
XX 16-JUL-1999; 99US-0144086P.
XX 19-JUL-1999; 99US-0144325P.
XX 19-JUL-1999; 99US-0144331P.
XX 19-JUL-1999; 99US-0144332P.
XX 19-JUL-1999; 99US-0144333P.
XX 19-JUL-1999; 99US-0144334P.
XX 19-JUL-1999; 99US-0144335P.
```


PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-01334256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-01351124P.
PR 21-MAY-1999; 99US-0135363P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 24-JUN-1999; 99US-0140823P.
PR 28-JUN-1999; 99US-0140991P.
PR 29-JUN-1999; 99US-0141287P.
PR 30-JUN-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147260P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149830P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.


```

PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 22-OCT-1999; 99US-0161404P.
PR 22-OCT-1999; 99US-0161405P.
PR 22-OCT-1999; 99US-0161406P.
PR 22-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 61.5%; Score 123; DB 3; Length 353;
Best Local Similarity 64.7%; Pred. No. 4.1e-09;
Matches 22; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 GFLGLMGSGIVSNLLKMGHTVTWNRTAEKCD 34
    ||||:||||: ||: ||||| |||
DB 65 GFLGIMGSPMAQNLIKAGCDVTWNRTKSKCD 98

RESULT 15
AAE36391
ID AAE36391 standard; protein; 290 AA.
XX AC AAE36391;
XX DT 26-JUN-2003 (first entry)
XX DE Tulip pistil gamma-hydroxybutyrate dehydrogenase homologue.
XX KW Alpha-methylene-gamma-butyrolactone; glutamate decarboxylase; herbicide;
XX KW enzyme; gamma-aminobutyrate aminotransferase; UDP-glucosyltransferase;
XX KW gamma-hydroxybutyrate dehydrogenase; tulipalin A; plant.
XX OS Tulip pistil.
XX PN WC2002101013-A2.
XX PD 19-DEC-2002.
XX PF 10-JUN-2002; 2002WO-US018230.
XX PR 08-JUN-2001; 2001US-0297198P.
XX PA (DUPO ) DU PONT DE NEMOURS & CO E I.
XX PI (PRAB/) PRABHU V.
XX PI Damude HG, Flint D, Prabhu V, Wang H;
XX WPI; 2003-201331/19.
XX DR N-PSDB; AAE55024.
XX PT Novel isolated nucleic acid fragment encoding a tuliposide A synthesizing
XX PT protein, useful for creating recombinant organisms that have the ability
XX PT to synthesize tulipalin A, tuliposide A or tuliposide A pathway
XX PT intermediates.
XX PS Claim 4; Page 110-111; 71pp; English.
XX CC The invention relates to genes encoding key enzymes in the biosynthesis
XX CC of alpha-methylene-gamma-butyrolactone (tulipalin A). Key enzymes include
XX CC glutamate decarboxylase, gamma-aminobutyrate aminotransferase, gamma-
XX CC hydroxybutyrate dehydrogenase and UDP-glucosyltransferase. The invention
XX CC is useful for producing tulipalin A or tuliposide A or its pathway
XX CC intermediates such as alpha-methylene-succinate semialdehyde, alpha-
XX CC methylene-gamma-aminobutyrate or alpha-methylene-gamma-hydroxybutyrate.
XX CC Tulipalin A sequences are used to immunise animals to produce polyclonal

CC or monoclonal antibodies with specificity for them or as targets to
CC facilitate design and/or identification of inhibitors of those enzymes
CC that may be useful as herbicides. The present sequence is Tulip pistil
CC gamma-hydroxybutyrate dehydrogenase homologue
XX SQ Sequence 290 AA;
Query Match 59.0%; Score 118; DB 6; Length 290;
Best Local Similarity 55.3%; Pred. No. 1.7e-08;
Matches 21; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 1 GFLGLMGSGIVSNLLKMGHTVTWNRTAEKCDLFIQ 38
    ||||:||||: ||: ||||| |||
DB 4 GFLGIMGKAMAVNLLRSGFRVTWNRTLSKCNELLE 41

RESULT 16
AAE19927
ID AAE19927 standard; protein; 360 AA.
XX AC AAE19927;
XX DT 18-JUN-2002 (first entry)
XX DE Corn 3-hydroxyisobutyrate dehydrogenase #2.
XX KW Branched chain amino acid degradation enzyme; HMG-CoA lyase;
XX KW 3-hydroxyisobutyrate dehydrogenase; 3-hydroxymethylglutaryl CoA;
XX KW isovaleryl-CoA dehydrogenase; transgenic plant; corn; EC 1.1.1.31.
XX OS Zea mays.
XX PN US6348339-B1.
XX PD 19-FEB-2002.
XX PF 29-JUL-1999; 99US-00364230.
XX PR 31-JUL-1998; 98US-0094990P.
XX PA (DUPO ) DU PONT DE NEMOURS & CO E I.
XX PI Cahoon RE, Hitz WD, Kinney AJ, Rafalski JA;
XX WPI; 2002-204621/26.
XX DR N-PSDB; AAD31748.
XX PT Isolated polynucleotide encoding 3-hydroxyisobutyrate dehydrogenase, used
XX PT to produce transgenic plants having an altered expression of the enzyme.
XX PS Claim 1; Col 27-30; 40pp; English.
XX CC The invention relates to nucleic acid fragments encoding branched chain
XX CC amino acid degradation enzymes. Particularly the invention relates to 3-
XX CC hydroxyisobutyrate dehydrogenase, 3-hydroxymethylglutaryl CoA (HMG-CoA)
XX CC lyase and isovaleryl-CoA dehydrogenase polypeptides and polynucleotides
XX CC derived from corn, rice and wheat. Sequences of the invention are used to
XX CC produce transgenic plants having an altered expression of the enzyme.
XX CC Polynucleotides of the invention can be used as probes for physical
XX CC mapping of genomes. The present sequence is corn 3-hydroxyisobutyrate
XX CC dehydrogenase (EC 1.1.1.31) from crln.pk0191.b10, p0076.cwhan08r and
XX CC p0118.chsbe01r clones
XX SQ Sequence 360 AA;
Query Match 58.5%; Score 117; DB 5; Length 360;
Best Local Similarity 63.6%; Pred. No. 3.1e-08;
Matches 21; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 GFLGLMGSGIVSNLLKMGHTVTWNRTAEKCD 33
    ||||:||||: ||: ||||| |||
DB 69 GFLGIMGKAMATNLLRHGFRVTWNRTLAKC 101

```

SO Sequence 292 AA;

Query Match 54.5%; Score 109; DB 6; Length 292;
Best Local Similarity 62.5%; Pred. No. 3.4e-07;
Matches 20; Conservative 6; Mismatches 6; Indels

Qy 1 GFGLGLMGSGIVSNLKMGHVTVVNRTAEK 32
 |||:|||:-:|||||||:-:||||
Dd 9 GFCTGVMGKGMIKNLLKGGYTVHVYNRTAEK 40

RESULT 18

ABU17331
ID ABU17331 standard; protein; 288 AA.

AA ABU17331:

19-JUN-2003 (first entry)

protein encoded by prokaryotic essential gene #2858.

XX **proliferative essential gene:** cell proliferation; drug design.

XX
acinetobacter baumannii.

XX WQ200277183-A2.

XX
PD 03-OCT-2002.

21-MAR-2002: 2002WO-US009107.

21-MAR-2001; 2001US-00815242.

PK 00-355-2001; 2001US-0342923P.
PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-0362699P.
PB 06-MAR-2002; 2002US-0362699P.

XX
D3 (EIT T-) ELITRA PHARM INC.

XX	Pt	Wang L,	Zamudio C,	Malone C,	Haselbeck R,	Ohlsen XL,	Zyskind JW;
		Carr GT,	Vamamoto R,	Forsyth RA,	Xu HH;

XX
DR WPT: 2003-029926/02.

XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT establish functional drug discovery programs.

XX
XX
XX
claim 25. SEO ID NO 45255: 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC

RESULT 17
ABU24159

AB024139
TN AB024159 standard; protein; 292 AA.

AC ABU24159:

19-JUN-2003 (first entry)

XX protein encoded by Prokaryotic essential gene #9686.

XX anti-oncogene; protooncogene; essential gene; cell proliferation; drug design.

XX
oc clocstriidum acetobutylicum.

XX PN WO200277183-A2.

XX
PD 03-OCT-2002.

21-MAR-2002: 2002WO-US009107.

21-MAR-2001: 2001US-00815242.

PR 25-OCT-2001: 2001US-0342923P.

PK 08-FEB-2002, 2002US-0362699P.
PP 06-MAR-2002: 2002US-0362699P.

XX
PA (ELIT-) ELITRA PHARM INC.

PJ					Zyskind JW;
XX					
PI	Wang L,	Zamudio C,	Hasselbeck R,	Ohlson KL,	
XI	wall G,	Carr GT,	Forsyth RA,	Xu HH;	
XX					

[illegible]

DR N-PSDB; ACA28029.

PT New antisense nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to

XX
claim 25. SEQ ID NO 52083. 1766bp. English:

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at fr-wipo.int/wmb/published/pct_sequences

33 XX

CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 288 AA;

Query Match 53.0%; Score 106; DB 6; Length 288;
 Best Local Similarity 55.6%; Pred. No. 9.2e-07;
 Matches 20; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 2 FLGLGMSGIVSNLLKMGHTVTVNRTAEKCDLFI 37
 DB 11 FLGLGMSGIVSNLLKMGHTVTVNRTAEKCDLFI 46

RESULT 19
 ID ADA36120 standard; protein; 299 AA.
 XX AC ADA36120;
 XX AC ADA36120;

DT 20-NOV-2003 (first entry)
 XX AC Acinetobacter baumannii protein #3281.
 DE Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
 KW Acinetobacter baumannii; plant biocontrol agent.
 KW Acinetobacter baumannii.

OS Acinetobacter baumannii.
 XX US6562958-B1.
 PN US6562958-B1.
 XX 13-MAY-2003.

XX 04-JUN-1999; 99US-00328352.
 XX 09-JUN-1998; 98US-0088701P.
 XX (GENO-) GENOME THERAPEUTICS CORP.

XX Breton G, Bush D;
 XX WPI; 2003-576092/54.
 DR N-PSDB; ADA31994.
 XX

PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
 PT for diagnosing a bacterial disease, as components of antibacterial
 PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
 PT plants.

PS Example; SEQ ID NO 7407; 328pp; English.
 XX

CC The invention relates to isolated Acinetobacter baumannii nucleic acids.
 CC The A. baumannii nucleic acids and polypeptides are useful as reagents
 CC for diagnosing a bacterial disease, as components of antibacterial
 CC vaccines, as targets for antibacterial drugs, to detect the presence of
 CC A. baumannii and other Acinetobacter species in a sample, in screening
 CC compounds for the ability to interfere with the A. baumannii life cycle
 CC or to inhibit A. baumannii infection, and as biocontrol agents for
 CC plants. The present sequence represents the amino acid sequence of an A.
 CC baumannii protein.
 XX

SQ Sequence 299 AA;

Query Match 53.0%; Score 106; DB 6; Length 299;
 Best Local Similarity 52.9%; Pred. No. 9.6e-07;
 Matches 18; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 1 FLGLGMSGIVSNLLKMGHTVTVNRTAEKCD 34
 DB 14 GFVGIGMGPMMANLLKMGHTVTVNRTAEKCD 47

RESULT 20

ABU17066
 ID ABU17066 standard; protein; 290 AA.
 XX AC ABU17066;
 XX AC ABU17066;

DT 19-JUN-2003 (first entry)
 XX DE Protein encoded by Prokaryotic essential gene #2593.
 XX DE Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX KW Acinetobacter baumannii.
 OS Acinetobacter baumannii.

PN WO200277183-A2.
 XX PD 03-OCT-2002.
 XX PF 21-MAR-2002; 2002WO-US009107.
 XX PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 DR N-PSDB; ACA20936.
 XX

PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX

PS Claim 25; SEQ ID NO 44990; 1766pp; English.
 XX

CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of

CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 290 AA;
Query Match 52.5%; Score 105; DB 6; Length 290;
Best Local Similarity 59.4%; Pred. No. 1.3e-06;
Matches 19; Conservative 4; Mismatches 9; Indels 0; Gaps 0;
Qy 1 FLGLGLMGSGIVSNLKGHTVTVMRTAEK 32
||:||||:||||:||||:||||:||||:|
Db 5 GFVGTGIMGMPMAMNLLKAGHQVAVNRTTSK 36

RESULT 21
ADA36852
ID ADA36852 standard; protein; 326 AA.
XX AC ADA36852;
XX DT 20-NOV-2003 (first entry)
XX DE Acinetobacter baumannii protein #4013.
XX DE Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
XX KW plant biocontrol agent.
XX OS Acinetobacter baumannii.
XX OS US6562958-B1.
XX FN 13-MAY-2003.
XX PF 04-JUN-1999; 99US-00328352.
XX PR 09-JUN-1998; 98US-0088701P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Breton G, Bush D;
XX WPI; 2003-576092/54.
XX DR N-PSDB; ADA32726.
XX PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
XX PT for diagnosing a bacterial disease, as components of antibacterial
XX PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
XX PT plants.
XX PS Example; SEQ ID NO 8139; 328pp; English.
XX CC The invention relates to isolated Acinetobacter baumannii nucleic acids.
XX CC The A. baumannii nucleic acids and polypeptides are useful as reagents
XX CC for diagnosing a bacterial disease, as components of antibacterial
XX CC vaccines, as targets for antibacterial drugs, to detect the presence of
XX CC A. baumannii and other Acinetobacter species in a sample, in screening
XX CC compounds for the ability to interfere with the A. baumannii life cycle
XX CC or to inhibit A. baumannii infection, and as biocontrol agents for
XX CC plants. The present sequence represents the amino acid sequence of an A.
XX CC baumannii protein.
XX SQ Sequence 326 AA;
Query Match 51.0%; Score 102; DB 6; Length 326;
Best Local Similarity 52.8%; Pred. No. 4.1e-06;
Matches 19; Conservative 5; Mismatches 12; Indels 0; Gaps 0;
Qy 2 FLGLGLMGSGIVSNLKGHTVTVMRTAEKDLFI 37
||:||||:||||:||||:||||:||||:|
Db 49 FLGLGLMGSRMASRLIQAGFQVAVNRTTSACELI 84

RESULT 22
AAG10425
ID AAG10425 standard; protein; 285 AA.
XX AC AAG10425;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 8741.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX FN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
XX PR 05-MAR-1999; 99US-0123180P.
XX PR 09-MAR-1999; 99US-0123548P.
XX PR 23-MAR-1999; 99US-0125788P.
XX PR 25-MAR-1999; 99US-0126264P.
XX PR 29-MAR-1999; 99US-0126785P.
XX PR 01-APR-1999; 99US-0127462P.
XX PR 06-APR-1999; 99US-0128234P.
XX PR 08-APR-1999; 99US-0128714P.
XX PR 16-APR-1999; 99US-0129845P.
XX PR 19-APR-1999; 99US-0130077P.
XX PR 21-APR-1999; 99US-0130449P.
XX PR 23-APR-1999; 99US-0130510P.
XX PR 28-APR-1999; 99US-0130891P.
XX PR 30-APR-1999; 99US-0131449P.
XX PR 30-APR-1999; 99US-0132048P.
XX PR 04-MAY-1999; 99US-0132407P.
XX PR 05-MAY-1999; 99US-0132484P.
XX PR 06-MAY-1999; 99US-0132486P.
XX PR 06-MAY-1999; 99US-0132486P.
XX PR 07-MAY-1999; 99US-0132863P.
XX PR 11-MAY-1999; 99US-0134256P.
XX PR 14-MAY-1999; 99US-0134218P.
XX PR 14-MAY-1999; 99US-0134219P.
XX PR 14-MAY-1999; 99US-0134221P.
XX PR 18-MAY-1999; 99US-0134370P.
XX PR 19-MAY-1999; 99US-0134768P.
XX PR 20-MAY-1999; 99US-0134941P.
XX PR 21-MAY-1999; 99US-0135124P.
XX PR 21-MAY-1999; 99US-0135353P.
XX PR 24-MAY-1999; 99US-0135629P.
XX PR 25-MAY-1999; 99US-0136021P.
XX PR 27-MAY-1999; 99US-0136392P.
XX PR 28-MAY-1999; 99US-0136782P.
XX PR 01-JUN-1999; 99US-0137222P.
XX PR 03-JUN-1999; 99US-0137222P.
XX PR 04-JUN-1999; 99US-0137528P.
XX PR 07-JUN-1999; 99US-0137502P.
XX PR 08-JUN-1999; 99US-0137724P.
XX PR 08-JUN-1999; 99US-0138094P.
XX PR 10-JUN-1999; 99US-0138540P.
XX PR 10-JUN-1999; 99US-0138847P.
XX PR 14-JUN-1999; 99US-0139119P.
XX PR 16-JUN-1999; 99US-0139452P.
XX PR 16-JUN-1999; 99US-0139453P.
XX PR 17-JUN-1999; 99US-0139492P.
XX PR 18-JUN-1999; 99US-0139454P.
XX PR 18-JUN-1999; 99US-0139455P.
XX PR 18-JUN-1999; 99US-0139456P.
XX PR 18-JUN-1999; 99US-0139457P.

Db 4 GFITGVMGSAVARHLLLEAGHEVAVYNRTRAKAD 37

19-JUN-2003 (first entry)
Protein encoded by Prokaryotic essential gene #15494.
Antisense; prokaryotic essential gene; cell proliferation; drug design.
Enterococcus faecium.
WO200277183-A2.
03-OCT-2002.
21-MAR-2002; 2002WO-US009107.
21-MAR-2001; 2001US-00815242.
06-SEP-2001; 2001US-00948993.
25-OCT-2001; 2001US-0342923P.
08-FEB-2002; 2002US-00072851.
06-MAR-2002; 2002US-0362699P.
(ELIT-) ELITRA PHARM INC.
Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen Kl, Zyskind JW;
Wail D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
WPI; 2003-029926/02.
N-PSDB; ACA33837.
New antisense nucleic acids, useful for identifying proteins or screening
for homologous nucleic acids required for cellular proliferation to
isolate candidate molecules for rational drug discovery programs.
Claim 25; SEQ ID NO 57891; 1766pp; English.
The invention relates to an isolated nucleic acid comprising any one of
the 6213 antisense sequences given in the specification where expression
of the nucleic acid inhibits proliferation of a cell. Also included are:
(1) a vector comprising a promoter operably linked to the nucleic acid
encoding a polypeptide whose expression is inhibited by the antisense
nucleic acid; (2) a host cell containing the vector; (3) an isolated
polypeptide or its fragment whose expression is inhibited by the
antisense nucleic acid; (4) an antibody capable of specifically binding
the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
proliferation or the activity of a gene in an operon required for
proliferation; (7) identifying a compound that influences the activity of
the gene product or that has an activity against a biological pathway
required for proliferation, or that inhibits cellular proliferation; (8)
identifying a gene required for cellular proliferation or the biological
pathway in which a proliferation-required gene or its gene product lies
or a gene on which the test compound that inhibits proliferation of an
organism acts; (9) manufacturing an antibiotic; (10) profiling a
compound's activity; (11) a culture comprising strains in which the gene
product is overexpressed or underexpressed; (12) determining the extent
to which each of the strains is present in a culture or collection of
strains; or (13) identifying the target of a compound that inhibits the
proliferation of an organism. The antisense nucleic acids are useful for
identifying proteins or screening for homologous nucleic acids required
for cellular proliferation to isolate candidate molecules for rational
drug discovery programs, or for screening homologous nucleic acids
required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
K. pneumoniae or *P. aeruginosa*. The present sequence is encoded by one of
the target prokaryotic essential genes. Note: The sequence data for this
patent did not form part of the printed specification, but was obtained
in electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences

Query Match 50.0%; Score 100; DB 6; Length 295;
Best Local Similarity 52.9%; Pred. No. 7e-06;
Matches 18; Conservative 6; Mismatches 10; Indels 0; Gaps 0;
XX
SQ Sequence 295 AA;
1 GFLGLGMSGIVSNLLKMGHTVTWNRTAEKCD 34

RESULT 24
ADC95482
ID ADC95482 standard; protein; 299 AA.
XX
AC ADC95482;
XX
DT 01-JAN-2004 (first entry)
XX
DE E. faecium protein sequence SEQ ID 5109.
XX
KW Vaccine; urinary tract infection; bacteraemia; endocarditis; wound;
abdominal-pelvic infection.
XX
OS Enterococcus faecium.
XX
PN US6583275-B1.
XX
PD 24-JUN-2003.
XX
PF 30-JUN-1998; 98US-00107532.
XX
PR 02-JUL-1997; 97US-0051571P.
PR 14-MAY-1998; 98US-0085598P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Doucette-Stamm LA, Bush D;
XX
DR WPI; 2003-799836/75.
DR N-PSDB; ADC91828.
XX
PT New isolated nucleic acid derived from Enterococcus faecium encoding an
Enterococcus faecium polypeptide useful for detection, prevention and
treatment of a pathological condition resulting from a bacterial
infection.
XX
PS Example 1; SEQ ID NO 5109; 243pp; English.
XX
CC The invention relates to an isolated nucleic acid derived from
Enterococcus faecium encoding an Enterococcus faecium polypeptide having
one of 10 fully defined sequences given in the (for comprising 40
sequential nucleotides chosen from any of the nucleic acids, its
complement or sequences hybridising to it). Also included are a
recombinant vector comprising the nucleic acid operably linked to
transcription regulatory element, a cell comprising the vector and a
single-stranded probe comprising the nucleic acid. The nucleic acids are
chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.
CC The nucleic acids is useful for diagnosing pathological conditions
resulting from *E. faecium* bacterial infection (e.g. urinary tract
infection, bacteraemia, endocarditis, wounds and abdominal-pelvic
infection) and for screening drugs such as agonists and antagonists. The
nucleic acid is useful for recombinant production of *Candida albicans* -
derived peptides or antisense polypeptides. Pharmaceutical compositions
and vaccines containing the nucleic acid are useful for preventing or
treating Enterococcus faecium infections. The present sequence represents
one if the disclosed *E. faecium* proteins.
XX
SQ Sequence 299 AA;
Query Match 50.0%; Score 100; DB 7; Length 299;
Best Local Similarity 52.9%; Pred. No. 7.1e-06;
Matches 18; Conservative 6; Mismatches 10; Indels 0; Gaps 0;
XX
SQ

Query Match 50.0%; Score 100; DB 6; Length 295;
Best Local Similarity 52.9%; Pred. No. 7e-06;
Matches 18; Conservative 6; Mismatches 10; Indels 0; Gaps 0;
XX
SQ Sequence 295 AA;
1 GFLGLGMSGIVSNLLKMGHTVTWNRTAEKCD 34

QY 1 GFLGLGMSGIVSNLLKMGHTVTWNRTAEKCD 34
DB 8 GFITGVMGSAVARHLLLEAGHEVAVYNRTRAKAD 41

QY 1 GFLGLGMSGIVSNLLKMGHTVTWNRTAEKCD 34

RESULT 25

RESULT 25

AAW27613

ID AAW27613 standard; protein; 492 AA.

XX AC AAW27613;

XX DT 22-DEC-1997 (first entry)

XX DE Brevibacterium flavum 6-phosphogluconate dehydrogenase.

XX KW Brevibacterium flavum; 6-phosphogluconate dehydrogenase;

XX KW recombinant production; coryneform; bacterium; bacteria.

XX OS Brevibacterium flavum.

XX PN JP09224662-A.

XX PD 02-SEP-1997.

XX PF 23-FEB-1996; 96JP-00036346.

XX PR 23-FEB-1996; 96JP-00036346.

XX FA (MITU) MITSUBISHI CHEM CORP.

XX DR WPI; 1997-484097/45.

XX DR N-FSDB; AAT86035.

PT 6-phosphogluconate dehydrogenase - can be recombinantly produced by transforming coryneform bacteria with DNA molecule encoding it.

PS Claim 1; Page 5-7; 8pp; Japanese.

CC The present sequence is the Brevibacterium flavum JM-233 6-phosphogluconate dehydrogenase (6-PD). The 6-PD can be recombinantly produced by transforming coryneform bacteria with the DNA molecule encoding it

XX SQ Sequence 492 AA;

Query Match

Best Local Similarity 46.5%; Score 93; DB 2; Length 492;

Matches 17; Conservative 45.9%; Pred. No. 0.00013;

Mismatches 7; Gaps 0;

Indels 0;

Gaps 0;

QY 1 GFLGILMGSGIVSNLLKMGHTVTVNRTAKCDLFI 37

Db 19 GVUGLAVMGSLNARNFARNGHTVAVNKRSTDKLI 55

RESULT 26

AAG06417

ID AAG06417 standard; protein; 158 AA.

XX AC AAG06417;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 3182.

XX KW Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-00301439.

XX PR 25-FEB-1999; 99US-0121825P.

XX PR 05-MAR-1999; 99US-0123180P.

XX PR 09-MAR-1999; 99US-0123548P.

23-MAR-1999; 99US-0125788P.
 25-MAR-1999; 99US-0126264P.
 29-MAR-1999; 99US-0126785P.
 01-APR-1999; 99US-0127462P.
 06-APR-1999; 99US-0128234P.
 08-APR-1999; 99US-0128714P.
 16-APR-1999; 99US-0129845P.
 19-APR-1999; 99US-0130077P.
 21-APR-1999; 99US-0130449P.
 23-APR-1999; 99US-0130510P.
 28-APR-1999; 99US-0130891P.
 30-APR-1999; 99US-0131449P.
 04-APR-1999; 99US-0132048P.
 04-MAY-1999; 99US-0132407P.
 05-MAY-1999; 99US-0132484P.
 06-MAY-1999; 99US-0132485P.
 06-MAY-1999; 99US-0132486P.
 07-MAY-1999; 99US-0132487P.
 11-MAY-1999; 99US-0132863P.
 14-MAY-1999; 99US-0134256P.
 14-MAY-1999; 99US-0134218P.
 14-MAY-1999; 99US-0134219P.
 14-MAY-1999; 99US-0134221P.
 18-MAY-1999; 99US-0134370P.
 18-MAY-1999; 99US-0134768P.
 19-MAY-1999; 99US-0134941P.
 20-MAY-1999; 99US-0135124P.
 21-MAY-1999; 99US-0135353P.
 24-MAY-1999; 99US-0135629P.
 25-MAY-1999; 99US-0136021P.
 27-MAY-1999; 99US-0136392P.
 28-MAY-1999; 99US-0136782P.
 01-JUN-1999; 99US-0137222P.
 03-JUN-1999; 99US-0137528P.
 04-JUN-1999; 99US-0137528P.
 07-JUN-1999; 99US-0137724P.
 08-JUN-1999; 99US-0138094P.
 10-JUN-1999; 99US-0138540P.
 10-JUN-1999; 99US-0138847P.
 14-JUN-1999; 99US-0139119P.
 16-JUN-1999; 99US-0139452P.
 16-JUN-1999; 99US-0139453P.
 17-JUN-1999; 99US-0139492P.
 18-JUN-1999; 99US-0139454P.
 18-JUN-1999; 99US-0139455P.
 18-JUN-1999; 99US-0139456P.
 18-JUN-1999; 99US-0139457P.
 18-JUN-1999; 99US-0139458P.
 18-JUN-1999; 99US-0139459P.
 18-JUN-1999; 99US-0139460P.
 18-JUN-1999; 99US-0139461P.
 18-JUN-1999; 99US-0139462P.
 18-JUN-1999; 99US-0139463P.
 18-JUN-1999; 99US-0139750P.
 18-JUN-1999; 99US-0139763P.
 21-JUN-1999; 99US-0139817P.
 22-JUN-1999; 99US-0139899P.
 23-JUN-1999; 99US-0140353P.
 23-JUN-1999; 99US-0140354P.
 24-JUN-1999; 99US-0140695P.
 28-JUN-1999; 99US-0140823P.
 29-JUN-1999; 99US-0140991P.
 30-JUN-1999; 99US-0141287P.
 01-JUL-1999; 99US-0141842P.
 02-JUL-1999; 99US-0142154P.
 02-JUL-1999; 99US-0142055P.
 06-JUL-1999; 99US-0142390P.
 08-JUL-1999; 99US-0142803P.
 09-JUL-1999; 99US-0142920P.
 12-JUL-1999; 99US-0142977P.
 13-JUL-1999; 99US-0143542P.
 14-JUL-1999; 99US-0143624P.
 15-JUL-1999; 99US-0144005P.

Thu Sep 16 09:36:50 2004

PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 08-MAY-1999; 99US-0132484P.
PR 08-MAY-1999; 99US-0132485P.
PR 08-MAY-1999; 99US-0132486P.
PR 08-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139452P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 28-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.

PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144844P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.

PR	22-JUL-1999;	99US-01450855Z
PR	22-JUL-1999;	99US-01450807P
PR	22-JUL-1999;	99US-01450808P
PR	22-JUL-1999;	99US-01451512P
PR	22-JUL-1999;	99US-01451514P
PR	22-JUL-1999;	99US-01451515Z
PR	22-JUL-1999;	99US-0145218P
PR	23-JUL-1999;	99US-0145218P
PR	23-JUL-1999;	99US-0145224P
PR	26-JUL-1999;	99US-0145276P
PR	27-JUL-1999;	99US-01455913P
PR	27-JUL-1999;	99US-0145918P
PR	27-JUL-1999;	99US-0145919P
PR	28-JUL-1999;	99US-01459511P
PR	28-JUL-1999;	99US-01459511P
PR	02-AUG-1999;	99US-0146386P
PR	02-AUG-1999;	99US-0146388P
PR	02-AUG-1999;	99US-0146388P
PR	03-AUG-1999;	99US-0147038P
PR	04-AUG-1999;	99US-0147204P
PR	04-AUG-1999;	99US-0147302P
PR	05-AUG-1999;	99US-0147192P
PR	05-AUG-1999;	99US-0147260P
PR	06-AUG-1999;	99US-0147303P
PR	06-AUG-1999;	99US-0147431P
PR	09-AUG-1999;	99US-0147493P
PR	09-AUG-1999;	99US-0147935P
PR	10-AUG-1999;	99US-0148171P
PR	11-AUG-1999;	99US-0148319P
PR	12-AUG-1999;	99US-0148341P
PR	13-AUG-1999;	99US-0148565P
PR	13-AUG-1999;	99US-0148684P
PR	16-AUG-1999;	99US-0149368P
PR	17-AUG-1999;	99US-0149157P
PR	18-AUG-1999;	99US-0149426P
PR	20-AUG-1999;	99US-0149722P
PR	20-AUG-1999;	99US-0149723P
PR	20-AUG-1999;	99US-0149829P
PR	23-AUG-1999;	99US-0149902P
PR	23-AUG-1999;	99US-0149930P
PR	25-AUG-1999;	99US-0150566P
PR	26-AUG-1999;	99US-0150884P
PR	27-AUG-1999;	99US-0151065P
PR	27-AUG-1999;	99US-0151066P
PR	27-AUG-1999;	99US-0151080P
PR	30-AUG-1999;	99US-0151303P
PR	31-AUG-1999;	99US-0151438P
PR	01-SEP-1999;	99US-0151930P
PR	07-SEP-1999;	99US-0152363P
PR	10-SEP-1999;	99US-0153070P
PR	13-SEP-1999;	99US-0153758P
PR	15-SEP-1999;	99US-0154018P
PR	16-SEP-1999;	99US-0154039P
PR	20-SEP-1999;	99US-0154779P
PR	23-SEP-1999;	99US-0155139P
PR	23-SEP-1999;	99US-0155486P
PR	24-SEP-1999;	99US-0155659P
PR	28-SEP-1999;	99US-0156458P
PR	29-SEP-1999;	99US-0156569P
PR	04-OCT-1999;	99US-0157117P
PR	05-OCT-1999;	99US-0157753P
PR	06-OCT-1999;	99US-0157865P
PR	07-OCT-1999;	99US-0158029P
PR	08-OCT-1999;	99US-0158232P
PR	12-OCT-1999;	99US-0158369P
PR	13-OCT-1999;	99US-0159293P
PR	13-OCT-1999;	99US-0159294P
PR	13-OCT-1999;	99US-0159330P
PR	14-OCT-1999;	99US-0159331P
PR	14-OCT-1999;	99US-0159637P
PR	14-OCT-1999;	99US-0159638P
PR	18-OCT-1999;	99US-0159584P
PR	21-OCT-1999;	99US-0160741P
PR	21-OCT-1999;	99US-0160767P

PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
Query Match 46.0%; Score 92; DB 3; Length 182;		
Best Local Similarity 45.9%; Pred. No. 5.5e-05;		
Matches 17; Conservative 8; Mismatches 12; Indels		
QY	1 GFLGLGMSGSIVNLLKMGCHVTVWNRPAEKCDLFI 37 :: : : :: :: }	
Dd	43 GWIGTGMGRSMCGHLIKAGYTTFVFNRTISKAOQLI 79 	
<hr/>		
RESULT 29		
AARG20956		
ID	AARG20956 standard; protein; 312 AA.	
XX	AC	AARG20956;
XT	AC	
DT	17-OCT-2000 (first entry)	
XX		
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 23337.	
XX		
KW	Protein identification; signal transduction pathway; metabo-	
KW	hybridisation assay; genetic mapping; gene expression cont-	
KW	termination sequence.	
XX		
OS	Arabidopsis thaliana.	
XX		
PN	EP1033405-A2.	
XX		
PD	06-SEP-2000.	
XX		
Pf	25-FEB-2000; 2000EP-00301439.	
XX		
PR	25-FEB-1999;	99US-012825P.
PR	05-MAR-1999;	99US-0123180P.
PR	09-MAR-1999;	99US-0123548P.
PR	23-MAR-1999;	99US-0125788P.
PR	25-MAR-1999;	99US-0126264P.
PR	29-MAR-1999;	99US-0126785P.
PR	01-APR-1999;	99US-0127462P.
PR	06-APR-1999;	99US-0128234P.
PR	08-APR-1999;	99US-0128714P.
PR	16-APR-1999;	99US-0129845P.
PR	19-APR-1999;	99US-0130077P.
PR	21-APR-1999;	99US-0130449P.
PR	23-APR-1999;	99US-0130510P.
PR	23-APR-1999;	99US-0130891P.
PR	28-APR-1999;	99US-0131449P.
PR	30-APR-1999;	99US-0132048P.
PR	30-APR-1999;	99US-0132407P.
PR	04-MAY-1999;	99US-0132484P.
PR	05-MAY-1999;	99US-0132485P.
PR	06-MAY-1999;	99US-0132486P.
PR	07-MAY-1999;	99US-0132487P.
PR	07-MAY-1999;	99US-0132863P.
PR	11-MAY-1999;	99US-0134256P.
PR	14-MAY-1999;	99US-0134218P.

PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 14-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-01351124P.
PR 21-MAY-1999; 99US-0135353P.
PR 21-MAY-1999; 99US-0135623P.
PR 24-MAY-1999; 99US-0136021P.
PR 25-MAY-1999; 99US-0136392P.
PR 27-MAY-1999; 99US-0136782P.
PR 28-MAY-1999; 99US-0137222P.
PR 01-JUN-1999; 99US-0137528P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139111P.
PR 14-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149923P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155653P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158023P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159323P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 22-OCT-1999; 99US-0160989P.

```
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 26-OCT-1999; 99US-0161362P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 46.0%; Score 92; DB 3; Length 312;
Best Local Similarity 45.9%; Pred. No. 0.00011;
Matches 17; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

Qy 1 GFLGLGLMGSGIVSNLLKMGHTVTIVNRTAEKCDLFI 37
Db 19 GWIGVGWGRSCWGLIKAGYIVTVFNRTISKAQTLI 55

RESULT 30
AAG20955
ID AAG20955 standard; protein; 334 AA.
XX
AC AAG20955;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 23336.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPL033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 26-OCT-1999; 99US-0161362P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139494P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140685P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
```

Search completed: September 16, 2004, 07:45:14
Job time : 62 secs